

QY 901 QEDRSSGTTTELHCVTDERNALRRSSAAHTSHNTYNFTKSENSNRCTCMPYAKLEYKRSS 960  
 Db 900 QEDRSSGTTTELHCVTDERNALRRSSAAHTSHNTYNFTKSENSNRCTCMPYAKLEYKRSS 959  
 QY 961 NDSLNSVSDCYGRGQKQKPSIESYSEDDSKFCSCYGOYPADIAHKTHSANHWDDNDGE 1020  
 Db 960 NDSLNSVSDGGRGQKQKPSIESYSEDDSKFCSCYGOYPADIAHKTHSANHWDDNDGE 1019  
 QY 1021 LDTPIYNSLYKDEQLNSGRQSPSQNERWARPKHIIIEDEIKQSEQRQSRNQSTTYPVYTE 1080  
 Db 1020 LDTPIYNSLYKDEQLNSGRQSPSQNERWARPKHIIIEDEIKQSEQRQSRNQSTTYPVYTE 1079  
 QY 1081 STDDXHLFPQHFQGOECVSPYRSRGANGSETNRVGSNHGINONVQSLOCEDDYEDDKP 1140  
 Db 1080 STDDXHLFPQHFQGOECVSPYRSRGANGSETNRVGSNHGINONVQSLOCEDDYEDDKP 1139  
 QY 1141 TNSERYSEEBEHEERPTNYSIKYNEEKHVDQPIDYSLKYATDIPSSQKQSPFSKS 1200  
 Db 1140 TNSERYSEEBEHEERPTNYSIKYNEEKHVDQPIDYSLKYATDIPSSQKQSPFSKS 1199  
 QY 1201 SSGQSKTEHMSSENTSTPSSNAKQNLHPSSAQSRSGOPQKAAATCKVSSINQETIQ 1260  
 Db 1200 SSGQSKTEHMSSENTSTPSSNAKQNLHPSSAQSRSGOPQKAAATCKVSSINQETIQ 1259  
 QY 1261 TYCVEDTPICFRCRCSLSLSAEDEIGCNQTTQADGANTLOIAEIKGIGTRSAEDPV 1320  
 Db 1260 TYCVEDTPICFRCRCSLSLSAEDEIGCNQTTQADGANTLOIAEIKGIGTRSAEDPV 1319  
 QY 1321 SEVPAVSQHPRTKSSRLQSSLSSESARHAKAVEFPSSGAKSPSKGAQTPKSPPEHYVOET 1380  
 Db 1320 SEVPAVSQHPRTKSSRLQSSLSSESARHAKAVEFPSSGAKSPSKGAQTPKSPPEHYVOET 1379  
 QY 1381 PLMFBSCTSVSLDSFESISASSVQSEPCSGMVGIIISPSLDPSQGTMPSPSKTTP 1440  
 Db 1380 PLMFBSCTSVSLDSFESISASSVQSEPCSGMVGIIISPSLDPSQGTMPSPSKTTP 1439  
 QY 1441 PPQTQATXREVPKNAKAPTAERESGPKQAAVNAAVQVQLPDAATLHFAFESTPDGF 1500  
 Db 1440 PPQTQATXREVPKNAKAPTAERESGPKQAAVNAAVQVQLPDAATLHFAFESTPDGF 1499  
 QY 1501 SCSSLSALSLEBPTQKDVLRIMPVOENONGNETESEQKESNENOEKAEKTIIDSE 1560  
 Db 1500 SCSSLSALSLEBPTQKDVLRIMPVOENONGNETESEQKESNENOEKAEKTIIDSE 1559  
 QY 1561 KDLDDSDDDDDIIELEECIIISAMPTKSRGKKPAQATASKLPPVARKPSQLPVVKLLPS 1620  
 Db 1560 KDLDDSDDDDDIIELEECIIISAMPTKSRGKKPAQATASKLPPVARKPSQLPVVKLLPS 1619  
 QY 1621 QNRLQPKHVSFTFGDDMPRVYCEGTPINFSTATSLSDLTIESPNELAAEGVGRGGAQ 1680  
 Db 1620 QNRLQPKHVSFTFGDDMPRVYCEGTPINFSTATSLSDLTIESPNELAAEGVGRGGAQ 1679  
 QY 1681 SGFEFKRDTIPTGRSTDEAOGKTSVITPELDDNKABEGDILAEICINSAMPKGSHPK 1740  
 Db 1680 SGFEFKRDTIPTGRSTDEAOGKTSVITPELDDNKABEGDILAEICINSAMPKGSHPK 1739  
 QY 1741 FRVKKINDQVQOASASSAPKQKQKKTSPVKPIQNTFYRVRKNADSKNNLN 1800  
 Db 1740 FRVKKINDQVQOASASSAPKQKQKKTSPVKPIQNTFYRVRKNADSKNNLN 1799  
 QY 1801 AERFVSDNKKQKQNLKQNSKDFNDKLPNNEDVRGSPAFDSPHHTPIEGTFCFSRND 1860  
 Db 1800 AERFVSDNKKQKQNLKQNSKDFNDKLPNNEDVRGSPAFDSPHHTPIEGTFCFSRND 1859  
 QY 1861 SLASLDFDDDDVDLSREKAEIRAKENKESAEKVTSHTELTSNQOQANKTQAIKQPINR 1920  
 Db 1860 SLASLDFDDDDVDLSREKAEIRAKENKESAEKVTSHTELTSNQOQANKTQAIKQPINR 1919  
 QY 1921 GQPKPILOKQSTFPQSSKDIIPDRGAATDEKLQNFATIENTPVCFSHNSSLSLSDIQENN 1980  
 Db 1920 GQPKPILOKQSTFPQSSKDIIPDRGAATDEKLQNFATIENTPVCFSHNSSLSLSDIQENN 1979

QY 1981 NKENEPKETEPPDPSQGBSPKQASGYAPKSPHEVEDTPVCFGRNSSLSLSDSEDDLLQ 2040  
 Db 1980 NKENEPKETEPPDPSQGBSPKQASGYAPKSPHEVEDTPVCFGRNSSLSLSDSEDDLLQ 2039  
 QY 2041 ECISAMPKPKKSRLLKQNEKXSPRNMGGIIGEDLTLKDIQRPDSEHGLSPDSENF 2100  
 Db 2040 ECISAMPKPKKSRLLKQNEKXSPRNMGGIIGEDLTLKDIQRPDSEHGLSPDSENF 2099  
 QY 2101 WKAIQEGANSIVSLHQAACCLSQASDSDSIISLSGSIISGSPFHLTPDOEKPF 2160  
 Db 2100 WKAIQEGANSIVSLHQAACCLSQASDSDSIISLSGSIISGSPFHLTPDOEKPF 2159  
 QY 2161 SNKGPRLLKPEGEKSTLTETKXIESEKIGKGVKYLITGKVRNSSEISGQKQPLQAN 2220  
 Db 2160 SNKGPRLLKPEGEKSTLTETKXIESEKIGKGVKYLITGKVRNSSEISGQKQPLQAN 2219  
 QY 2221 MPISIRGRTHIHPGVNRNSSSTSPVSKGPPKLTPTASKSPSEGOATTSPPRAKSPVKS 2280  
 Db 2220 MPISIRGRTHIHPGVNRNSSSTSPVSKGPPKLTPTASKSPSEGOATTSPPRAKSPVKS 2279  
 QY 2281 ELSPVARTSOIGSSKAPSRSGRSDTPGRPAQOPLSRPIQSPGRNSISPGRNGISPPN 2340  
 Db 2280 ELSPVARTSOIGSSKAPSRSGRSDTPGRPAQOPLSRPIQSPGRNSISPGRNGISPPN 2339  
 QY 2341 KLSQLPRTSPSTASTKSSGSGKMYTSPGRQMSQNLTKQTGLSKNASSIPRSESASKG 2400  
 Db 2340 KLSQLPRTSPSTASTKSSGSGKMYTSPGRQMSQNLTKQTGLSKNASSIPRSESASKG 2399  
 QY 2401 LNQMNGANGANKVELSRMSTKSSGSESDRSPVLRVROSTPIKEAPSTLARKLEESA 2460  
 Db 2400 LNQMNGANGANKVELSRMSTKSSGSESDRSPVLRVROSTPIKEAPSTLARKLEESA 2459  
 QY 2461 SFESLSFSSRPASPTRQAOQTPVLSPLDMSLSTHSSVQAGGWRKLPPLNLSPTIEYNDG 2520  
 Db 2460 SFESLSFSSRPASPTRQAOQTPVLSPLDMSLSTHSSVQAGGWRKLPPLNLSPTIEYNDG 2519  
 QY 2521 RPAKRHDIARSHSESPRLPINRSGTWKRHSKHSSSLPRVSTWRTGSSSSILSASES 2580  
 Db 2520 RPAKRHDIARSHSESPRLPINRSGTWKRHSKHSSSLPRVSTWRTGSSSSILSASES 2579  
 QY 2581 SEKAKSDEKHVNSISGTQKSKENOVSAKGTWRKI KENEFSPNTNSTQTVSSGATNGAES 2640  
 Db 2580 SEKAKSDEKHVNSISGTQKSKENOVSAKGTWRKI KENEFSPNTNSTQTVSSGATNGAES 2639  
 QY 2641 KTLIYQMAPAVSKTEDVWVRIEDCPINNPRSGRSPGTGNTPPVIDSVSEKANPNIKDSKDN 2700  
 Db 2640 KTLIYQMAPAVSKTEDVWVRIEDCPINNPRSGRSPGTGNTPPVIDSVSEKANPNIKDSKDN 2699  
 QY 2701 QAKQNVGNGSVPMETVGLNLTSTFIQVDAPDQKGTETIKQGNPNPVVSETNESPIVERT 2760  
 Db 2700 QAKQNVGNGSVPMETVGLNLTSTFIQVDAPDQKGTETIKQGNPNPVVSETNESPIVERT 2759  
 QY 2761 PFSSSSSSKHSPPSGTVAARVTPFNYPSPRKSADSTSAAPSQIPTPVNNNTKKRDSKT 2820  
 Db 2760 PFSSSSSSKHSPPSGTVAARVTPFNYPSPRKSADSTSAAPSQIPTPVNNNTKKRDSKT 2819  
 QY 2821 DSTSSSGTQSPKRHSGLYVTSV 2843  
 Db 2820 DSTSSSGTQSPKRHSGLYVTSV 2842

Search completed: August 25, 2004, 17:14:55  
 Job time : 198.5 secs

**This Page Blank (uspio)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 25, 2004, 17:23:12 ; Search time 190.5 Seconds

(without alignments)  
4895.247 Million cell updates/sec

Title: US-09-442-489F-2

Perfect score: 14575

Sequence: 1 MAASVYDQLKQVFAKWMEN.....ESSGTQSPKRRHSGSYLVTSV 2843

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1297172 seqs, 314612898 residues

Total number of hits satisfying chosen parameters: 1297172

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA.\*

- 1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pcp.\*
- 2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pcp.\*
- 3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pcp.\*
- 4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pcp.\*
- 5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pcp.\*
- 6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pcp.\*
- 7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pcp.\*
- 8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pcp.\*
- 9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pcp.\*
- 10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pcp.\*
- 11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pcp.\*
- 12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pcp.\*
- 13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pcp.\*
- 14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pcp.\*
- 15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pcp.\*
- 16: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pcp.\*
- 17: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pcp.\*
- 18: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pcp.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	14533	99.7	2843	9	US-09-987-482-1
2	14533	99.7	2843	12	US-10-392-113-21
3	14526.5	99.7	2844	12	US-10-267-502-370
4	14524	99.7	2843	16	US-10-408-765A-1970
5	14506	99.5	2843	8	US-08-681-219-32
6	14506	99.5	2843	11	US-09-230-111C-30
7	14506	99.5	2843	14	US-10-092-138-30
8	13148	90.2	2845	12	US-10-267-502-372
9	4719	32.4	912	9	US-09-987-482-2
10	3889	26.7	767	9	US-09-987-482-3
11	3508.5	24.1	2274	12	US-10-267-502-373
12	3414.5	23.4	2303	12	US-10-267-502-371
13	1763	12.1	2417	12	US-10-267-502-368
14	1640	11.3	325	12	US-09-915-307-5
15	1474	10.1	1633	14	US-10-029-386-33090

Query Match 99.7%; Score 14533; DB 9; Length 2843;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 2836; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Qy	1	MAASVYDQLKQVFAKWMENLNHLEEDNSNHLTKLETSANNKEVYKQLOQSSIDEAM	60
Db	1	MAASVYDQLKQVFAKWMENLNHLEEDNSNHLTKLETSANNKEVYKQLOQSSIDEAM	60
Qy	61	ASSGQIDLLERLKLNDSSNFPQVKLSRSMLSYSGREGSVSSRSGECSVPMGSPFR	120
Db	61	ASSGQIDLLERLKLNDSSNFPQVKLSRSMLSYSGREGSVSSRSGECSVPMGSPFR	120
Qy	121	RGFVNGRESGYLLEEKERSLLADLDKEEKEKDWYQAQLNLTTRIDSLPTENFSL	180
Db	121	RGFVNGRESGYLLEEKERSLLADLDKEEKEKDWYQAQLNLTTRIDSLPTENFSL	180
Qy	191	QTDLTROLEVEARQIRVAMEEQIGTCDQMEKRAQRRIATQIQEKKDILRLTROLLSOAT	240

ALIGNMENTS

RESULT 1

US-09-987-482-1  
; Sequence 1, Application US/09987482  
; Publication No. US20020184656A1  
; GENERAL INFORMATION:  
; APPLICANT: BHANDARI, POONAM  
; APPLICANT: SHASHIDHARA, L.S.  
; TITLE OF INVENTION: IN VIVO ASSAY SYSTEM FOR SCREENING AND VALIDATION OF  
; TITLE OF INVENTION: DRUGS AND OTHER SUBSTANCES  
; FILE REFERENCE: 056859-0134  
; CURRENT APPLICATION NUMBER: US/09/987,482  
; CURRENT FILING DATE: 2002-03-21  
; NUMBER OF SEQ ID NOS: 3  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 1  
; LENGTH: 2843  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-987-482-1

Sequence 369, Appl  
Sequence 12713, A  
Sequence 2231, Ap  
Sequence 43924, A  
Sequence 4, Appli  
Sequence 49, Appl  
Sequence 104793,  
Sequence 142, App  
Sequence 70580, A  
Sequence 23, Appl  
Sequence 5, Appli  
Sequence 5784, Ap  
Sequence 143, App  
Sequence 2287, Ap  
Sequence 73678, A  
Sequence 144, App  
Sequence 83, Appl  
Sequence 71033, A  
Sequence 34248, A  
Sequence 11, Appl  
Sequence 8, Appli  
Sequence 5, Appli  
Sequence 71235, A  
Sequence 62, Appl  
Sequence 81, Appl  
Sequence 359, App  
Sequence 12, Appl  
Sequence 8, Appli

Db 181 QTDNTRQLEYEARQIRVAMEEQJGTQCDMEKRAQRRIARIQOIEKDILIRIOLLOSOAT 240  
Qy 241 EAERSSQNKHETGSHDAERQNEGQVGEINMATSNGQGGTTRMDHETASVLSSSTHSA 300  
Db 241 EAERSSQNKHETGSHDAERQNEGQVGEINMATSNGQGGTTRMDHETASVLSSSTHSA 300  
Qy 301 PRRLTSHLGTKVEMYLLNGLGTHDXDMSRTLAMSSQDSCINMRQGCPLLIQLL 360  
Db 301 PRRLTSHLGTKVEMYLLNGLGTHDXDMSRTLAMSSQDSCINMRQGCPLLIQLL 360  
Qy 361 HGNDKDSVLLGNSRGSKEARAPASAAAHNIHSDPDDKRRREIRVLHLEQIRAYCETC 420  
Db 361 HGNDKDSVLLGNSRGSKEARAPASAAAHNIHSDPDDKRRREIRVLHLEQIRAYCETC 420  
Qy 421 WEWQEAHEPGMDQKNMPAPVEHQICPACVCLMKLSFDEEHRHAMNELGGLQAI AELLQ 480  
Db 421 WEWQEAHEPGMDQKNMPAPVEHQICPACVCLMKLSFDEEHRHAMNELGGLQAI AELLQ 480  
Qy 481 VDCEMYGLTNDHYSITLRRYAGMALTNLTFGDVANKATILCSMGCCRVAUVAQILKSEEDL 540  
Db 481 VDCEMYGLTNDHYSITLRRYAGMALTNLTFGDVANKATILCSMGCCRVAUVAQILKSEEDL 540  
Qy 541 QOVVASVLRNLNLSRADVNSKKTLEVGSVKALMECALEVKESTLKSVALNLSAHC 600  
Db 541 QOVVASVLRNLNLSRADVNSKKTLEVGSVKALMECALEVKESTLKSVALNLSAHC 600  
Qy 601 ENKADICAVDGAFLVGTLYTYSQNTLAIIESGGGILRWSSLIATNEDHROILRENN 660  
Db 601 ENKADICAVDGAFLVGTLYTYSQNTLAIIESGGGILRWSSLIATNEDHROILRENN 660  
Qy 661 CLQTLLOHLKSHSITIVSNACGLTNLSARNPKDQALWDMGAVMLKNLHSHKHWMIA 720  
Db 661 CLQTLLOHLKSHSITIVSNACGLTNLSARNPKDQALWDMGAVMLKNLHSHKHWMIA 720  
Qy 721 GSAAALRNLMANRPKADANIMSPGSLPSLHVRRKQKALEAEADQHLSETFDNDNLS 780  
Db 721 GSAAALRNLMANRPKADANIMSPGSLPSLHVRRKQKALEAEADQHLSETFDNDNLS 780  
Qy 781 PKASHRSKORHQSILYGDYVDTNRHDDNDSFNFTGNMTVLSPYNTTIVLPSSSSRGS 840  
Db 781 PKASHRSKORHQSILYGDYVDTNRHDDNDSFNFTGNMTVLSPYNTTIVLPSSSSRGS 840  
Qy 841 LDSSRSKRSLEMERGIGLGNHPATENPGTSSKRGLOISTTAAQIAKVMESVAIHTS 900  
Db 841 LDSSRSKRSLEMERGIGLGNHPATENPGTSSKRGLOISTTAAQIAKVMESVAIHTS 900  
Qy 901 QEDRSSGTTLHCVTDERNALRRSSAAHSHNTYFTKSENRRCTSPYAKLEYKRS 960  
Db 901 QEDRSSGTTLHCVTDERNALRRSSAAHSHNTYFTKSENRRCTSPYAKLEYKRS 960  
Qy 961 NDSLNSVSSNDGYKRGQMKPSIESYSEDDKFCVGYQYPADLAHKIHSANHMDNDGE 1020  
Db 961 NDSLNSVSSNDGYKRGQMKPSIESYSEDDKFCVGYQYPADLAHKIHSANHMDNDGE 1020  
Qy 1021 LDTPTNLSKYDBQLNSGRQSPONERWAPKHIIIEDEIKQSEQRNQSTTYPVYTE 1080  
Db 1021 LDTPTNLSKYDBQLNSGRQSPONERWAPKHIIIEDEIKQSEQRNQSTTYPVYTE 1080  
Qy 1081 STDDKHLKFQPHFQOCVCPYRSRGANGSETNVRVSGNHGINQVQSILCOEDDYEKKP 1140  
Db 1081 STDDKHLKFQPHFQOCVCPYRSRGANGSETNVRVSGNHGINQVQSILCOEDDYEKKP 1140  
Qy 1141 TNSYERYSEBEQHEEBRPTNYSIKYNEEKRRVDQPIDYSLKVATDIPSSQKQSFSSKS 1200  
Db 1141 TNSYERYSEBEQHEEBRPTNYSIKYNEEKRRVDQPIDYSLKVATDIPSSQKQSFSSKS 1200  
Qy 1201 SSGOSSKTEHMSSESSENTSPSSNAKQONLHPSSAQSRSGQPKAATCKVSSINQETIQ 1260  
Db 1201 SSGOSSKTEHMSSESSENTSPSSNAKQONLHPSSAQSRSGQPKAATCKVSSINQETIQ 1260  
Qy 1261 TYCVEDTPICFSRCSLSLSAABDEICGNQTTQEADPSANTLQIABIKGIGTRSAEDPV 1320  
Db 1261 TYCVEDTPICFSRCSLSLSAABDEICGNQTTQEADPSANTLQIABIKGIGTRSAEDPV 1320

Qy 1321 SEVPAVSQHPRTKSRLOGSSLSSESARHKAVERPPSGAKSPSKGAOTPKPSPPEHYQET 1380  
Db 1321 SEVPAVSQHPRTKSRLOGSSLSSESARHKAVERPPSGAKSPSKGAOTPKPSPPEHYQET 1380  
Qy 1381 PLMPFRCTSVSSLSDFESRSIASSVQSEPCSGMVSGIISPSDLDPSPGQOTWPPRSKTPP 1440  
Db 1381 PLMPFRCTSVSSLSDFESRSIASSVQSEPCSGMVSGIISPSDLDPSPGQOTWPPRSKTPP 1440  
Qy 1441 PPQTAQTKREVKNKAPTAEKRESGPKQAANAARVQVLPADATLLHFATESTPDGF 1500  
Db 1441 PPQTAQTKREVKNKAPTAEKRESGPKQAANAARVQVLPADATLLHFATESTPDGF 1500  
Qy 1501 SCSSLSALSLEPFIQKVELRIMPVQENDNGNETESEQPKESNENQEKAEKTI DSE 1560  
Db 1501 SCSSLSALSLEPFIQKVELRIMPVQENDNGNETESEQPKESNENQEKAEKTI DSE 1560  
Qy 1561 KDLDDSDDDDIIELEECIIISAMPTKSRKKGKPAQTASKLPPPVARKPSOLPVYKLLPS 1620  
Db 1561 KDLDDSDDDDIIELEECIIISAMPTKSRKKGKPAQTASKLPPPVARKPSOLPVYKLLPS 1620  
Qy 1621 QNRLQPKXVSFTPGDDMPRVYCVEGTPIINFSTATSLDLTIESPPNELAAGEGVGGAQ 1680  
Db 1621 QNRLQPKXVSFTPGDDMPRVYCVEGTPIINFSTATSLDLTIESPPNELAAGEGVGGAQ 1680  
Qy 1681 SGEPFKRDTIIPTEGRSTDEAOGGKTSSVTIPELDONKAEEDGIIAECINSAMPKGSHP 1740  
Db 1681 SGEPFKRDTIIPTEGRSTDEAOGGKTSSVTIPELDONKAEEDGIIAECINSAMPKGSHP 1740  
Qy 1741 FRVKIMDQVQOASASSAPNKNQLDGKKKXPTSPVKPIPQNTYRTRVRKNADSKNIN 1800  
Db 1741 FRVKIMDQVQOASASSAPNKNQLDGKKKXPTSPVKPIPQNTYRTRVRKNADSKNIN 1800  
Qy 1801 ABRVFSNDKSKQNKLKNSKDFNDKLPNNEDRVGSPAFDSPHYTYPIEGTPTCFSRND 1860  
Db 1801 ABRVFSNDKSKQNKLKNSKDFNDKLPNNEDRVGSPAFDSPHYTYPIEGTPTCFSRND 1860  
Qy 1861 SLSSLDLDDDDVDLSREKAEELRKAKENKESAKVTSHTELTSNOOSANKTOAIKQPINR 1920  
Db 1861 SLSSLDLDDDDVDLSREKAEELRKAKENKESAKVTSHTELTSNOOSANKTOAIKQPINR 1920  
Qy 1921 GQPKPILQKQSTFPQSSKOIPDRGAATDEKLQNPATIENTPVCFSHNSLSLSLSDIDENN 1980  
Db 1921 GQPKPILQKQSTFPQSSKOIPDRGAATDEKLQNPATIENTPVCFSHNSLSLSLSDIDENN 1980  
Qy 1981 NKENEPIKETPPSQCEPSPKQASGVAPKSFHVEDTPVCFSRNSSLSLSDIDEDLLQ 2040  
Db 1981 NKENEPIKETPPSQCEPSPKQASGVAPKSFHVEDTPVCFSRNSSLSLSDIDEDLLQ 2040  
Qy 2041 ECISAMPKKKPSRLKGDNEKHSFRNMGGLGDETLDLKDIQRPDSEHGLSPDSNFD 2100  
Db 2041 ECISAMPKKKPSRLKGDNEKHSFRNMGGLGDETLDLKDIQRPDSEHGLSPDSNFD 2100  
Qy 2101 WKATOEGANSIVSLHQAAAACLSROASDSDSILSKGIGLSGSPFHLTPDOEKPF 2160  
Db 2101 WKATOEGANSIVSLHQAAAACLSROASDSDSILSKGIGLSGSPFHLTPDOEKPF 2160  
Qy 2161 SNKGPRILKPEKSTLETTKIETSESKGIGKKGVYKSLITGKVRNSGEISQMKQPLQAN 2220  
Db 2161 SNKGPRILKPEKSTLETTKIETSESKGIGKKGVYKSLITGKVRNSGEISQMKQPLQAN 2220  
Qy 2221 MPSISRGTMTHIFGVNRSSSSTSPVKKGPPLKTPASKSPSECOATTTPRGAKPSVKS 2280  
Db 2221 MPSISRGTMTHIFGVNRSSSSTSPVKKGPPLKTPASKSPSECOATTTPRGAKPSVKS 2280  
Qy 2281 ELSPVARTSOIGGSSKAPSRSGSRDSTPSRPAQOPLSRPIQSPGRNSISPGRNGISPPN 2340  
Db 2281 ELSPVARTSOIGGSSKAPSRSGSRDSTPSRPAQOPLSRPIQSPGRNSISPGRNGISPPN 2340  
Qy 2341 KLSQLPRTSSPTASTKSSGSGKSVYTPGRMQOQNLTKQTGLSKVASSIPRSESASKG 2400  
Db 2341 KLSQLPRTSSPTASTKSSGSGKSVYTPGRMQOQNLTKQTGLSKVASSIPRSESASKG 2400



```
QY 2401 LNMNNGANKKVELSRMSSKSGESDRSFRPVLVRQSTFIKEAPSPILRRKLEBSA 2460
DB 2401 LNMNNGANKKVELSRMSSKSGESDRSFRPVLVRQSTFIKEAPSPILRRKLEBSA 2460
QY 2461 SPESLSPSRPASPTRSQATPVLSPSLPDMSLSTHSSVQAGWKLPPNLSPTTIEYNDG 2520
DB 2461 SPESLSPSRPASPTRSQATPVLSPSLPDMSLSTHSSVQAGWKLPPNLSPTTIEYNDG 2520
QY 2521 RPAKHDIARSHSESPSRPIPNRSCTWKEHSHKSSSLPRVSTWRTGSSSILSASSES 2580
DB 2521 RPAKHDIARSHSESPSRPIPNRSCTWKEHSHKSSSLPRVSTWRTGSSSILSASSES 2580
QY 2581 SEKAKSEDEKHVNSISGTKQSKENOVSAKGTWRKIKENEFSPTNSTSTQTVSSGATNGAES 2640
DB 2581 SEKAKSEDEKHVNSISGTKQSKENOVSAKGTWRKIKENEFSPTNSTSTQTVSSGATNGAES 2640
QY 2641 KTLIYQMAPAVSKTEDVWVRIEDCPINNPRSGRSPGTNTPPVIDSVSEKANPINIKSDKN 2700
DB 2641 KTLIYQMAPAVSKTEDVWVRIEDCPINNPRSGRSPGTNTPPVIDSVSEKANPINIKSDKN 2700
QY 2701 QAKQNVGSGVPMRTVGLNRLTSTFQVDAPQKGTETKPGQNNPVVSETNESPIVERT 2760
DB 2701 QAKQNVGSGVPMRTVGLNRLTSTFQVDAPQKGTETKPGQNNPVVSETNESPIVERT 2760
QY 2761 PFSSSSSSKHSSPSGTVAAARVTPFNYPNPSRKSSADTSARPSQIPTPVNNNTKKRDSKT 2820
DB 2761 PFSSSSSSKHSSPSGTVAAARVTPFNYPNPSRKSSADTSARPSQIPTPVNNNTKKRDSKT 2820
QY 2821 DSTESSGTOSPKRHSGSYLVTSV 2843
DB 2821 DSTESSGTOSPKRHSGSYLVTSV 2843

RESULT 2
US-10-392-113-21
; Sequence 21, Application US/10392113
; Publication No. US2003024993A1
; GENERAL INFORMATION:
; APPLICANT: Land, Hartmut
; TITLE OF INVENTION: COMPOSITIONS THAT INHIBIT PROLIFERATION
; FILE OF INVENTION: OF CANCER CELLS
; FILE REFERENCE: 21108.00053
; CURRENT APPLICATION NUMBER: US/10/392,113
; CURRENT FILING DATE: 2003-03-17
; PRIOR APPLICATION NUMBER: 60/365,078
; PRIOR FILING DATE: 2002-03-15
; PRIOR APPLICATION NUMBER: PCT/US01/32127
; PRIOR FILING DATE: 2001-10-15
; PRIOR APPLICATION NUMBER: 60/239,705
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 21
; LENGTH: 2843
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:/Note =
; OTHER INFORMATION: Synthetic Construct
US-10-392-113-21

Query Match 99.7%; Score 14533; DB 12; Length 2843;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2836; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 MAAAYDQLLKQVEALKMENSLNRQLEEDNSNHLTKLETEASNKMEVLKQLQGSIEDEAM 60
DB 1 MAAAYDQLLKQVEALKMENSLNRQLEEDNSNHLTKLETEASNKMEVLKQLQGSIEDEAM 60
QY 61 ASSGQDILLERLKEINLDSSNPFQVLRKMSLRSYGSREGSVSSRSGCSVPVPMGSPFR 120
DB 61 ASSGQDILLERLKEINLDSSNPFQVLRKMSLRSYGSREGSVSSRSGCSVPVPMGSPFR 120
```

```
QY 121 RGFVNGSRESTGYLLELEKERSILLADLDKEEKEKDWYAOQLNLTFRIDSLPLTENFSL 180
DB 121 RGFVNGSRESTGYLLELEKERSILLADLDKEEKEKDWYAOQLNLTFRIDSLPLTENFSL 180
QY 181 QTDLTRQLQYEAREQIRVAMEEQIGTCCDMKEKAQRRIARIQOITEKDILRIRQLQSQAT 240
DB 181 QTDLTRQLQYEAREQIRVAMEEQIGTCCDMKEKAQRRIARIQOITEKDILRIRQLQSQAT 240
QY 241 EAESSONKHETGSHDAERQNEGGVGEINMAYSGNGCGSTTRMDHETASVLSSSSTHSA 300
DB 241 EAESSONKHETGSHDAERQNEGGVGEINMAYSGNGCGSTTRMDHETASVLSSSSTHSA 300
QY 301 PRRLTSHLGTKEVMYVLSMLGTHDKDDMSRTLLAMSSQDSCISMROGCGCLPLLIQLL 360
DB 301 PRRLTSHLGTKEVMYVLSMLGTHDKDDMSRTLLAMSSQDSCISMROGCGCLPLLIQLL 360
QY 361 HGNDKDSVLLGNSRSGSEARASAAALHNTIHSQDDDKRGRREIRVUHLLEQIRAYCETC 420
DB 361 HGNDKDSVLLGNSRSGSEARASAAALHNTIHSQDDDKRGRREIRVUHLLEQIRAYCETC 420
QY 421 WEMOEAEHPGMDQDKNMPAPVEHOICPAVCVLMLKLSFDEEHRHAMNELGGLQIAELLQ 480
DB 421 WEMOEAEHPGMDQDKNMPAPVEHOICPAVCVLMLKLSFDEEHRHAMNELGGLQIAELLQ 480
QY 481 VDCMYGLTNDHYISITLRRYAGMALTNLTFGDVANKATLCMKGCMRALVAQLKSESED 540
DB 481 VDCMYGLTNDHYISITLRRYAGMALTNLTFGDVANKATLCMKGCMRALVAQLKSESED 540
QY 541 QOVTASVLRNLSWRADVNSKKTIREVGSVKALMECALEVKESTLKVLSALMNLASHCT 600
DB 541 QOVTASVLRNLSWRADVNSKKTIREVGSVKALMECALEVKESTLKVLSALMNLASHCT 600
QY 601 ENKADICAVDGALAFVLGTLTYSQTNTLAIIESGGGILRNVSLSIATNEDHROILRENN 660
DB 601 ENKADICAVDGALAFVLGTLTYSQTNTLAIIESGGGILRNVSLSIATNEDHROILRENN 660
QY 661 CLQTLLOHLKXSLTIIVSNACGTLWNLARNPKDOEALWDMGAVSMKXNLHSHKXMIAM 720
DB 661 CLQTLLOHLKXSLTIIVSNACGTLWNLARNPKDOEALWDMGAVSMKXNLHSHKXMIAM 720
QY 721 GSAALRNLMANRPKAKYKDNIMSPGSSLPDLVRKOKALEAEALDAOHLSETFNDNLNLS 780
DB 721 GSAALRNLMANRPKAKYKDNIMSPGSSLPDLVRKOKALEAEALDAOHLSETFNDNLNLS 780
QY 781 PKASHRSKQRHKQSLYGDYVFDNRRHDDNRSDNFTGNTMTVLSPLYLNTTVLPSSSSSRGS 840
DB 781 PKASHRSKQRHKQSLYGDYVFDNRRHDDNRSDNFTGNTMTVLSPLYLNTTVLPSSSSSRGS 840
QY 841 LDSSRSKDBSLERERGIIGLVNHPATENPCTSSKRGLOJSTTAAQAKWEEVSAIHTS 900
DB 841 LDSSRSKDBSLERERGIIGLVNHPATENPCTSSKRGLOJSTTAAQAKWEEVSAIHTS 900
QY 901 QEDRSSGTTTELHCVTDERNALRRSSAAHSTNFTKSENSNRITCSPYAKLEYKRSS 960
DB 901 QEDRSSGTTTELHCVTDERNALRRSSAAHSTNFTKSENSNRITCSPYAKLEYKRSS 960
QY 961 NDSLSNVSSNDGCKRGOMKPSIESYEDDESKECSYGOVPADLAHKHSHANHMDNDGE 1020
DB 961 NDSLSNVSSNDGCKRGOMKPSIESYEDDESKECSYGOVPADLAHKHSHANHMDNDGE 1020
QY 1021 LDTPTNYSKYSDEQLNSGRQSPQNERWARPPKHIIIEDEIKQSEQRQSRNQSTTYPVYTE 1080
DB 1021 LDTPTNYSKYSDEQLNSGRQSPQNERWARPPKHIIIEDEIKQSEQRQSRNQSTTYPVYTE 1080
QY 1081 STDDKHLKQFPHGQOCQVSPYRSGANGSETNRVGNHGINQVNSOSLQOEDDYEDDKP 1140
DB 1081 STDDKHLKQFPHGQOCQVSPYRSGANGSETNRVGNHGINQVNSOSLQOEDDYEDDKP 1140
QY 1141 TNSERYSEBQHEEBERPTNYSIKYNEEKHVDQPIDYSLKYATDTPSSQKQSFSPKS 1200
DB 1141 TNSERYSEBQHEEBERPTNYSIKYNEEKHVDQPIDYSLKYATDTPSSQKQSFSPKS 1200
```

QY 1201 SSGSSKTEHMSSESTSTPSSNAKQNLQHPSSAQSRSQPOKAATCKVSSINQETIQ 1260  
DB 1201 SSGSSKTEHMSSESTSTPSSNAKQNLQHPSSAQSRSQPOKAATCKVSSINQETIQ 1260  
QY 1261 TYCVEDTPICFSRGSSLSLSAABDEICGNOTTOEASANTLQIAEIKGKIGTRSAEDPV 1320  
DB 1261 TYCVEDTPICFSRGSSLSLSAABDEICGNOTTOEASANTLQIAEIKGKIGTRSAEDPV 1320  
QY 1321 SEVPAVSHPRTKSRRLQSSLSSESARHKAVERPPSGAKSPSKCAOTPKSPPEHYVOET 1380  
DB 1321 SEVPAVSHPRTKSRRLQSSLSSESARHKAVERPPSGAKSPSKCAOTPKSPPEHYVOET 1380  
QY 1381 PLMFSTRCTSVSLDSFESRSIASVQSEPCSGMVSGIISPSDLPDSFGQTMPFSRSKTPP 1440  
DB 1381 PLMFSTRCTSVSLDSFESRSIASVQSEPCSGMVSGIISPSDLPDSFGQTMPFSRSKTPP 1440  
QY 1441 PPPQTAQTKREVKNKAPTAKRESGPKQAANAAVORVQVLPDADTLLHFAIESTPDGF 1500  
DB 1441 PPPQTAQTKREVKNKAPTAKRESGPKQAANAAVORVQVLPDADTLLHFAIESTPDGF 1500  
QY 1501 SCSSLSALSDEPFIQXDELVRIMPVPOENDNGNETESQPKESNENQKEAEKTIIDSE 1560  
DB 1501 SCSSLSALSDEPFIQXDELVRIMPVPOENDNGNETESQPKESNENQKEAEKTIIDSE 1560  
QY 1561 KDLDDSDDDDIIEILBECIISAMPTKSRGKKAQATASKLPPVARKPSQLPVYKLLPS 1620  
DB 1561 KDLDDSDDDDIIEILBECIISAMPTKSRGKKAQATASKLPPVARKPSQLPVYKLLPS 1620  
QY 1621 QNRQPQKHVSFTPGDDMPRYVCVEGTPIINFSTATSLSDLTIESPPNELAAGEVGRGAQ 1680  
DB 1621 QNRQPQKHVSFTPGDDMPRYVCVEGTPIINFSTATSLSDLTIESPPNELAAGEVGRGAQ 1680  
QY 1681 SGEFEKRTIITEGRSTDEAOGKTSVTPPELDDNKAEGDILAEICINSAMPKGSHP 1740  
DB 1681 SGEFEKRTIITEGRSTDEAOGKTSVTPPELDDNKAEGDILAEICINSAMPKGSHP 1740  
QY 1741 FRVKIMQVOQAASASAPNNKQLDGKKKPTSPVKPIQONTYRTRVRKNADSKNNLN 1800  
DB 1741 FRVKIMQVOQAASASAPNNKQLDGKKKPTSPVKPIQONTYRTRVRKNADSKNNLN 1800  
QY 1801 AERFSDNKSKQNLKNNKDFNDKLPNNEDRVGSGFAFDSPHHYTPIEGTYPYCFSRND 1860  
DB 1801 AERFSDNKSKQNLKNNKDFNDKLPNNEDRVGSGFAFDSPHHYTPIEGTYPYCFSRND 1860  
QY 1861 SLSSLDPDDDDVLSREKAEILKAKENKESAKVTSHTELTSNCOOSANKTOAQKQPINR 1920  
DB 1861 SLSSLDPDDDDVLSREKAEILKAKENKESAKVTSHTELTSNCOOSANKTOAQKQPINR 1920  
QY 1921 GQPKPILQKSTFPQSSXDIIDRGAATDEKLQNFPAIENTPVCFSHNSLSLSDIDQENN 1980  
DB 1921 GQPKPILQKSTFPQSSXDIIDRGAATDEKLQNFPAIENTPVCFSHNSLSLSDIDQENN 1980  
QY 1981 NKEKEPIKETPPSQSGPFPQASGAPKSFHVEDTPVCFSRNSSLSLSDIDQENLLQ 2040  
DB 1981 NKEKEPIKETPPSQSGPFPQASGAPKSFHVEDTPVCFSRNSSLSLSDIDQENLLQ 2040  
QY 2041 ECISAMPKKKKPSRLKGDNEKHSRPNMGGLIGEDLTLDLKDIOQRPDSEHGLSPDSNFD 2100  
DB 2041 ECISAMPKKKKPSRLKGDNEKHSRPNMGGLIGEDLTLDLKDIOQRPDSEHGLSPDSNFD 2100  
QY 2101 WKATQEGANSIVSSLHQAAAAACLSRQASSDSLSLKSIGLSGSPFHTPDQEEKPF 2160  
DB 2101 WKATQEGANSIVSSLHQAAAAACLSRQASSDSLSLKSIGLSGSPFHTPDQEEKPF 2160  
QY 2161 SNKGPRILKPEKSTLETFTKTESSEKGIKGGKVVYKSLITGKVRNSBEISGQMKQPLQAN 2220  
DB 2161 SNKGPRILKPEKSTLETFTKTESSEKGIKGGKVVYKSLITGKVRNSBEISGQMKQPLQAN 2220  
QY 2221 MPFSRGRMTIHIPGVNRSSSSTSPVSKGPPKLTTPASKSPSEGCATTATTPRGAKPVS 2280  
DB 2221 MPFSRGRMTIHIPGVNRSSSSTSPVSKGPPKLTTPASKSPSEGCATTATTPRGAKPVS 2280  
QY 2281 ELSPVARQTSQIGSSSKAPSRSGSRDSTPSPRPAQQLSRPIQSPGRNSISPGRNGISPPN 2340

DB 2281 ELSPVARQTSQIGSSSKAPSRSGSRDSTPSPRPAQQLSRPIQSPGRNSISPGRNGISPPN 2340  
QY 2341 KLSQIPRTSSPSTASTKSSGSKMSYTSFGQMSQOQLTKOTGLSKNASSIIPRESASKG 2400  
DB 2341 KLSQIPRTSSPSTASTKSSGSKMSYTSFGQMSQOQLTKOTGLSKNASSIIPRESASKG 2400  
QY 2401 LNQMNGANGANKVLSRMSSTKSSGESDESEREPVLVRQSTFIKEAPSPTLRKLLESA 2460  
DB 2401 LNQMNGANGANKVLSRMSSTKSSGESDESEREPVLVRQSTFIKEAPSPTLRKLLESA 2460  
QY 2461 SFESLSPSRPASPTRSOAOTPVLSPLPDMSLSTHSSVQAGWRKLPNLSPTIEYNDG 2520  
DB 2461 SFESLSPSRPASPTRSOAOTPVLSPLPDMSLSTHSSVQAGWRKLPNLSPTIEYNDG 2520  
QY 2521 RPAKRDHARSHSESPSRLPINRSGTWKREHSKHSSSLPRVSTWRTGSSSSILSASSES 2580  
DB 2521 RPAKRDHARSHSESPSRLPINRSGTWKREHSKHSSSLPRVSTWRTGSSSSILSASSES 2580  
QY 2581 SEKAKSEDEKHVNSISGTSQKQKNOVSAGTWRKI KENEFPTNSTSTQTVSSGATNGAES 2640  
DB 2581 SEKAKSEDEKHVNSISGTSQKQKNOVSAGTWRKI KENEFPTNSTSTQTVSSGATNGAES 2640  
QY 2641 KTLIYQMAPAVSKTDDVVRIEDCPINNPRGSRPTGNTPPVIDSVSEKANPNIKDSKN 2700  
DB 2641 KTLIYQMAPAVSKTDDVVRIEDCPINNPRGSRPTGNTPPVIDSVSEKANPNIKDSKN 2700  
QY 2701 QAKQNVGSGVPMRTVGLNRLTSTFIQVDAPOKQTEIKPGQNNPVPVSETNESPIVERT 2760  
DB 2701 QAKQNVGSGVPMRTVGLNRLTSTFIQVDAPOKQTEIKPGQNNPVPVSETNESPIVERT 2760  
QY 2761 PSSSSSSKHSGSPCTVAARVTPPNYNSPRKSSADSTSARPSQIPTVNNNTKRDSE 2820  
DB 2761 PSSSSSSKHSGSPCTVAARVTPPNYNSPRKSSADSTSARPSQIPTVNNNTKRDSE 2820  
QY 2821 DSTESSGTSQPKRHSGLVITSV 2843  
DB 2821 DSTESSGTSQPKRHSGLVITSV 2843  
RESULT 3  
US-10-267-502-370  
; Sequence 370, Application US/10267502  
; Publication No. US20040071700A1  
; GENERAL INFORMATION:  
; APPLICANT: Kim, Jaeseob  
; TITLE OF INVENTION: Obesity Linked Genes  
; FILE REFERENCE: LSD-07416  
; CURRENT APPLICATION NUMBER: US/10/267,502  
; CURRENT FILING DATE: 2003-01-27  
; NUMBER OF SEQ ID NOS: 439  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 370  
; LENGTH: 2844  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-267-502-370  
Query Match 99.7%; Score 14526.5; DB 12; Length 2844;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 2837; Conservative 1; Mismatches 5; Indels 1; Gaps 1;  
QY 1 MAASVDQLLKQVEALKMKNENLROLEDNSNHLTKLETSANMKVYLKQLQGSIEDAM 60  
DB 1 MAASVDQLLKQVEALKMKNENLROLEDNSNHLTKLETSANMKVYLKQLQGSIEDAM 60  
QY 61 ASSGQIDILLERIKELNLDSSNPPGVKLSKMSLRSGVSGSRSCEGSPVPMGSPFR 120  
DB 61 ASSGQIDILLERIKELNLDSSNPPGVKLSKMSLRSGVSGSRSCEGSPVPMGSPFR 120  
QY 121 RGCVNGSRESTGYLEBELEKERSLLADLDKBEKMDYTAQLQNLTKRIDSLPTENFSL 180

Db 121 RGFVNGSRSTGYLEELEKERSLILADLDKEEKEKDYAQLQNLTKRIDSLPLTENFSL 180  
 QY 181 QTDLTRRQLEYEARQIRVAMEEQIGTQODMEKRAORRIARIQIQTEKIDILRIQLLOQAT 240  
 Db 181 QTDLTRRQLEYEARQIRVAMEEQIGTQODMEKRAORRIARIQIQTEKIDILRIQLLOQAT 240  
 QY 241 EAEKSSQNKHETGSHDAERQNGOGVGEINWATSGNGOGSTTRMDHETASVLSSSSTHSA 300  
 Db 241 EAEKSSQNKHETGSHDAERQNGOGVGEINWATSGNGOGSTTRMDHETASVLSSSSTHSA 300  
 QY 301 PRRLTSHLGTKEVWYVLSLLMLGTHDKDDMSRTLLAMSSQDSCISMQRGCLPLLIQLL 360  
 Db 301 PRRLTSHLGTKEVWYVLSLLMLGTHDKDDMSRTLLAMSSQDSCISMQRGCLPLLIQLL 360  
 QY 361 HGNDKDSVLLGNSRGSKEARARASALHNIITHSQDDKSGREIRVJLHLEOIRAYCETC 420  
 Db 361 HGNDKDSVLLGNSRGSKEARARASALHNIITHSQDDKSGREIRVJLHLEOIRAYCETC 420  
 QY 421 WEMQEAHPGMDQDNPNPAPVEHQCIPAVCVLMKLSFDEEHRHAMELGLQIAIELLQ 480  
 Db 421 WEMQEAHPGMDQDNPNPAPVEHQCIPAVCVLMKLSFDEEHRHAMELGLQIAIELLQ 480  
 QY 481 VDCENYGLTNDHYSITLARYAGMALTNLTFGDVANKATLCSMKGCMRALVAQLKSEEDL 540  
 Db 481 VDCENYGLTNDHYSITLARYAGMALTNLTFGDVANKATLCSMKGCMRALVAQLKSEEDL 540  
 QY 541 QOVIASVLRNLSWRADVANSKTLREVGSVKALMECALVKEESTLKSVLISALMNLASHCT 600  
 Db 541 QOVIASVLRNLSWRADVANSKTLREVGSVKALMECALVKEESTLKSVLISALMNLASHCT 600  
 QY 601 ENKADICAVDGAFLVGLTVRSQNTNLATIESGGGILRVNVSLLIANNEHROILRENN 660  
 Db 601 ENKADICAVDGAFLVGLTVRSQNTNLATIESGGGILRVNVSLLIANNEHROILRENN 660  
 QY 661 CIQTLILHLKSHSLTIVSNACGTLANLSARNPKDOEALWMDGAVSMKLNLIHSHKMIAM 720  
 Db 661 CIQTLILHLKSHSLTIVSNACGTLANLSARNPKDOEALWMDGAVSMKLNLIHSHKMIAM 720  
 QY 721 GSAALRLNLMANRPARYKADANTWSPGSSILPSLHVVRKQKALEAELDAQHLSFTFDNIDLS 780  
 Db 721 GSAALRLNLMANRPARYKADANTWSPGSSILPSLHVVRKQKALEAELDAQHLSFTFDNIDLS 780  
 QY 781 PKASHRSKORHQSLYGDFVFTNDRHDDNRSDFNTGNMTVLSPLYNTTIVLPSSSSRGS 840  
 Db 781 PKASHRSKORHQSLYGDFVFTNDRHDDNRSDFNTGNMTVLSPLYNTTIVLPSSSSRGS 840  
 QY 841 LQSSRSKDRSLERERIGLGNVHPATENPGTSSKRGLOISTTAAQIAKVMEEVSAIHTS 900  
 Db 841 LQSSRSKDRSLERERIGLGNVHPATENPGTSSKRGLOISTTAAQIAKVMEEVSAIHTS 900  
 QY 901 QEDRSSTTELHCVTDERNALRRSSAHTHNTYNTFTKSENSNRTCSMPYAKLEYKRS 960  
 Db 901 QEDRSSTTELHCVTDERNALRRSSAHTHNTYNTFTKSENSNRTCSMPYAKLEYKRS 960  
 QY 961 NDLSNSVSSNDGKQKQKPSIESYEDDESKEFCYGOYPADLAHKIHSANHMDNDGE 1020  
 Db 961 NDLSNSVSSNDGKQKQKPSIESYEDDESKEFCYGOYPADLAHKIHSANHMDNDGE 1020  
 QY 1021 LQTPINYSKYSDQNLNSRQSPQNERWARPKHIEDEIKOSEQRQSRNQSTTVPVYTE 1080  
 Db 1021 LQTPINYSKYSDQNLNSRQSPQNERWARPKHIEDEIKOSEQRQSRNQSTTVPVYTE 1080  
 QY 1081 STDDKHLKQPHFGQOECPYRSGANGETNRVGSNHNQNTYSOSLCOEDDYDDKP 1140  
 Db 1081 STDDKHLKQPHFGQOECPYRSGANGETNRVGSNHNQNTYSOSLCOEDDYDDKP 1140  
 QY 1141 TNYSERVSEEOHEEBERPTNYSIKYNEEKRHVDQPIDYSLKYATDIPSSQKSFSSKS 1200  
 Db 1141 TNYSERVSEEOHEEBERPTNYSIKYNEEKRHVDQPIDYSLKYATDIPSSQKSFSSKS 1200  
 QY 1201 SSQGSKTEHMSSSSENTSTPSSNAKRONQLHPSSAQSRSGQPKAATCKVSSINQETIQ 1260  
 Db 1201 SSQGSKTEHMSSSSENTSTPSSNAKRONQLHPSSAQSRSGQPKAATCKVSSINQETIQ 1260

QY 1261 TYCVEDTPICFRCSLSSLSAEDEIGCNQVTTQBADSSANTLQIAIEIKGKIGTRSAEDPV 1320  
 Db 1261 TYCVEDTPICFRCSLSSLSAEDEIGCNQVTTQBADSSANTLQIAIEIKGKIGTRSAEDPV 1320  
 QY 1321 SEVP-AVSOHPRTKSRLOQSSLSSESAHKAVEPPSGAKSPSKGAQTPKPPPHYQOE 1379  
 Db 1321 SEVPSSVHSTLTETKSRLOQSSLSSESAHKAVEPPSGAKSPSKGAQTPKPPPHYQOE 1380  
 QY 1380 TPLMFRCCTSVSSLSDFESRSTASSVQSPCQSGWVGIISPSDLDPSPCQTMPPPSKTP 1439  
 Db 1381 TPLMFRCCTSVSSLSDFESRSTASSVQSPCQSGWVGIISPSDLDPSPCQTMPPPSKTP 1440  
 QY 1440 PPPQTAQTKREVPKNAKPTAEKRESGPKQAAVNAAVQVQVLPDADTLHLHFATESTPDG 1499  
 Db 1441 PPPQTAQTKREVPKNAKPTAEKRESGPKQAAVNAAVQVQVLPDADTLHLHFATESTPDG 1500  
 QY 1500 FSCSSLSLSALDEPFIQKDVLRIMPVQENDNGNETESEOPKESNENQEKAEKTIIDS 1559  
 Db 1501 FSCSSLSLSALDEPFIQKDVLRIMPVQENDNGNETESEOPKESNENQEKAEKTIIDS 1560  
 QY 1560 EKDLDDSDDDDEIIELEECIISAMPTKSRKGGKPAQATASKLPPPVARKPSQLPVYKLLP 1619  
 Db 1561 EKDLDDSDDDDEIIELEECIISAMPTKSRKGGKPAQATASKLPPPVARKPSQLPVYKLLP 1620  
 QY 1620 SQNRLOPQKHVSFTPGDDMPRVYCVGTPINFSTATISLSDLTIESPPNELAAGEVYRGA 1679  
 Db 1621 SQNRLOPQKHVSFTPGDDMPRVYCVGTPINFSTATISLSDLTIESPPNELAAGEVYRGA 1680  
 QY 1680 QSGFEKRDITPIEGHSTDEAOGKTSSTVITPELDNKAEBEGDILAEACINSAMPKCKSHK 1739  
 Db 1681 QSGFEKRDITPIEGHSTDEAOGKTSSTVITPELDNKAEBEGDILAEACINSAMPKCKSHK 1740  
 QY 1740 PFRVKIMDQVQASASSAPNKNQLDGKKKTPSPVKPIQONTYRTRVRKNADSKNNL 1799  
 Db 1741 PFRVKIMDQVQASASSAPNKNQLDGKKKTPSPVKPIQONTYRTRVRKNADSKNNL 1800  
 QY 1800 NAEVFSNDKSKKQNLKNNSKDFNDKLPNNEDRVGSPAFDPSFHHVTPTEGTPYCFSRN 1859  
 Db 1801 NAEVFSNDKSKKQNLKNNSKDFNDKLPNNEDRVGSPAFDPSFHHVTPTEGTPYCFSRN 1860  
 QY 1860 DLSLSLDFDDDDVDLSREKAEELKAKENKESAKVTSHTELTNSQOSANKTQAIKQPIN 1919  
 Db 1861 DLSLSLDFDDDDVDLSREKAEELKAKENKESAKVTSHTELTNSQOSANKTQAIKQPIN 1920  
 QY 1920 RGQPKPILOKQSTPPQSSKDIPOGAATDEKLNFAIENTPVCFSHNSLSLSDIDQEN 1979  
 Db 1921 RGQPKPILOKQSTPPQSSKDIPOGAATDEKLNFAIENTPVCFSHNSLSLSDIDQEN 1980  
 QY 1980 NKNENEPKETEPPDSQCEPSKPOASGYAPKSFHVEDTPVCFRNSLSLSIDSEDDL 2039  
 Db 1981 NKNENEPKETEPPDSQCEPSKPOASGYAPKSFHVEDTPVCFRNSLSLSIDSEDDL 2040  
 QY 2040 QECISSAMPKCKPSRLKGDNEKHSPRNMGILGEDTLDLKQIORDPDSEHGLSPDSENF 2099  
 Db 2041 QECISSAMPKCKPSRLKGDNEKHSPRNMGILGEDTLDLKQIORDPDSEHGLSPDSENF 2100  
 QY 2100 DWKAIQEGANSIVSLHCAAAAACLSQAQSDSDSILSLKSGISLGSFPFHLTPQOEKPF 2159  
 Db 2101 DWKAIQEGANSIVSLHCAAAAACLSQAQSDSDSILSLKSGISLGSFPFHLTPQOEKPF 2160  
 QY 2160 TSNKGPRLKPEGEKSTLTETKIESEKIGKGVYKSLITGKVRNSEISGOMKQPLQA 2219  
 Db 2161 TSNKGPRLKPEGEKSTLTETKIESEKIGKGVYKSLITGKVRNSEISGOMKQPLQA 2220  
 QY 2220 NMPISIRGRTHHTPGVRNNSSTSPVSKGPPKTPASKSPSEGTATTPSPRCAKPSVK 2279  
 Db 2221 NMPISIRGRTHHTPGVRNNSSTSPVSKGPPKTPASKSPSEGTATTPSPRCAKPSVK 2280  
 QY 2280 SELSPVARQTSQIGGSSKAPSRSGRSDSTPSRPAQQLSRPIQSPGRNISIPGRNGISPP 2339  
 Db 2281 SELSPVARQTSQIGGSSKAPSRSGRSDSTPSRPAQQLSRPIQSPGRNISIPGRNGISPP 2340

```
QY 2340 NKLSOLPRTSSPSTASTKSSGSGKMSYTPSGQMSQQNLTKQTLGSKNASSIPRESASK 2399
Db 2341 NKLSOLPRTSSPSTASTKSSGSGKMSYTPSGQMSQQNLTKQTLGSKNASSIPRESASK 2400
QY 2400 GLNQMNNGANGKVKELSRMSTKSSGSESRSRPRVLRQSTFKEAPSPTRKLEES 2459
Db 2401 GLNQMNNGANGKVKELSRMSTKSSGSESRSRPRVLRQSTFKEAPSPTRKLEES 2460
QY 2460 ASFSLSPPSPAPTRSQACTPVLSPSLPDMSLSTHSSVQAGGWRKLPNLSPTIYND 2519
Db 2461 ASFSLSPPSPAPTRSQACTPVLSPSLPDMSLSTHSSVQAGGWRKLPNLSPTIYND 2520
QY 2520 GRPAKRHDIAHSHSESRLPINRSGTWKREHSHSSSLPRVSTWRTGSSSILSASSE 2579
Db 2521 GRPAKRHDIAHSHSESRLPINRSGTWKREHSHSSSLPRVSTWRTGSSSILSASSE 2580
QY 2580 SSEKASEDEKHVNSISGTKQSKENQVSAGTWRKIKENEFSPNTNSTQTVSSGATNGAE 2639
Db 2581 SSEKASEDEKHVNSISGTKQSKENQVSAGTWRKIKENEFSPNTNSTQTVSSGATNGAE 2640
QY 2640 SKTLIYQWAPAVSKTEDVWRIEDCPINPRSGRSPGTGNTPPVIDSVSEKANPNIKDSKD 2699
Db 2641 SKTLIYQWAPAVSKTEDVWRIEDCPINPRSGRSPGTGNTPPVIDSVSEKANPNIKDSKD 2700
QY 2700 NOAKQNVGNSVPMRTVGLNRLTSFIQVADPOKGTETKPGQNNPVPVSETNESPIVER 2759
Db 2701 NOAKQNVGNSVPMRTVGLNRLTSFIQVADPOKGTETKPGQNNPVPVSETNESPIVER 2760
QY 2760 TPFSSSSSSKHSPPSGTVAARVTPFNPNPSPKSSADSTGARPSPQIPTPVNNNTKKRDSK 2819
Db 2761 TPFSSSSSSKHSPPSGTVAARVTPFNPNPSPKSSADSTGARPSPQIPTPVNNNTKKRDSK 2820
QY 2820 TDSTESSGTSQPKRHSGSYLVTSV 2843
Db 2821 TDSTESSGTSQPKRHSGSYLVTSV 2844

RESULT 4
US-10-408-765A-1970
; Sequence 1970, Application US/10408765A
; Publication No. US20040101874A1
; GENERAL INFORMATION:
; APPLICANT: Ghosh, Soumitra S.
; APPLICANT: Fahy, Eoin D.
; APPLICANT: Zhang, Bing
; APPLICANT: Gibson, Bradford W.
; APPLICANT: Taylor, Steven W.
; APPLICANT: Glenn, Gary M.
; APPLICANT: Warnock, Dale E.
; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
; FILE REFERENCE: 660088.465
; CURRENT APPLICATION NUMBER: US/10/408,765A
; NUMBER OF SEQ ID NOS: 3077
; SOFTWARE: PASCSEQ for Windows Version 4.0
; SEQ ID NO 1970
; LENGTH: 2843
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-408-765A-1970

Query Match 99.7%; Score 14524; DB 16; Length 2843;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 2835; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 1 MAAASYDQLLKQVEALKWMSNLRQLEDNNSHLTKLETEASNMEVLKQLOGSIEDAM 60
Db 1 MAAASYDQLLKQVEALKWMSNLRQLEDNNSHLTKLETEASNMEVLKQLOGSIEDAM 60
QY 61 ASSGQIDLLERLKEINLDSNPPGVKLRKMSLRVSGSREGSVSRSGECSPPVPMGSPFR 120
Db 61 ASSGQIDLLERLKEINLDSNPPGVKLRKMSLRVSGSREGSVSRSGECSPPVPMGSPFR 120
```

```
QY 121 RGFVNGSRESTGYLBELEKERSLLADLDKBEKEDWYQAOLQNLTKRDISLPTENFSL 180
Db 121 RGFVNGSRESTGYLBELEKERSLLADLDKBEKEDWYQAOLQNLTKRDISLPTENFSL 180
QY 181 QTDLTRRQLEYEARQIRVAMEEQGLCTQDMMEKRAQRRIARIQOIEKDIRLIRQLQSQAT 240
Db 181 QTDLTRRQLEYEARQIRVAMEEQGLCTQDMMEKRAQRRIARIQOIEKDIRLIRQLQSQAT 240
QY 241 EAERSQNKHETGSHDAERQNGEGGVGEINWATSGOGGSTRMDHETASTVLSSSSTHSA 300
Db 241 EAERSQNKHETGSHDAERQNGEGGVGEINWATSGOGGSTRMDHETASTVLSSSSTHSA 300
QY 301 PRRLTSHLGTKVEMYVYLLSMLGTHDKDDMSRTLLAMSSSDSCISMRQSCCLPILLQLL 360
Db 301 PRRLTSHLGTKVEMYVYLLSMLGTHDKDDMSRTLLAMSSSDSCISMRQSCCLPILLQLL 360
QY 361 HGNDKDSVILGNRSGSKARARASALHNIITHSQDDDKRGRREIRVHLLEQIRAYCETC 420
Db 361 HGNDKDSVILGNRSGSKARARASALHNIITHSQDDDKRGRREIRVHLLEQIRAYCETC 420
QY 421 MEWQEAHEPFGMDQDKNPMPPAVEHQICPAVCVLMKLSFDEHHRHAMELGGLQIAELLQ 480
Db 421 MEWQEAHEPFGMDQDKNPMPPAVEHQICPAVCVLMKLSFDEHHRHAMELGGLQIAELLQ 480
QY 481 VDCEMYGLTNDHYSITLRYAGMALTNLTFQGVANKATLCSMKGCMRALVAQLKSEEDL 540
Db 481 VDCEMYGLTNDHYSITLRYAGMALTNLTFQGVANKATLCSMKGCMRALVAQLKSEEDL 540
QY 541 QOVTASVLRNLSRADVNSKTLREVGSVKALMECALEVKKESTLKSVLALNLSAHCT 600
Db 541 QOVTASVLRNLSRADVNSKTLREVGSVKALMECALEVKKESTLKSVLALNLSAHCT 600
QY 601 ENKADICAVDGAFLVGTLYRSQTNLTAIESGGGILRVSSLIATNEDHRLRENN 660
Db 601 ENKADICAVDGAFLVGTLYRSQTNLTAIESGGGILRVSSLIATNEDHRLRENN 660
QY 661 CLQTLQHLKSHSLTIYSNACGLTNLSARNPKDOEALWDMGAVSMKNTLHSHKMIAM 720
Db 661 CLQTLQHLKSHSLTIYSNACGLTNLSARNPKDOEALWDMGAVSMKNTLHSHKMIAM 720
QY 721 GSAALRNLMANRPAPKNDANIMSPGSSLPSSLHVRKQKALEAEIDAQHLSETFDNIDLS 780
Db 721 GSAALRNLMANRPAPKNDANIMSPGSSLPSSLHVRKQKALEAEIDAQHLSETFDNIDLS 780
QY 781 PKASHRSKQRHKOSLYGDYVFDNTRHDNRSDFNTGNVTLSPLNTTVLPSSSSRGS 840
Db 781 PKASHRSKQRHKOSLYGDYVFDNTRHDNRSDFNTGNVTLSPLNTTVLPSSSSRGS 840
QY 841 LDSRSERKDRSLERERIGLGNVHPATENPCTSSKRGLOISTTAAQIAKYWEVSAIHTS 900
Db 841 LDSRSERKDRSLERERIGLGNVHPATENPCTSSKRGLOISTTAAQIAKYWEVSAIHTS 900
QY 901 QEDRSSGSTTTLHCVTDERNALRRSSAAHSTHNTYNTKSENSNRTCSMPYAKLEYKRSS 960
Db 901 QEDRSSGSTTTLHCVTDERNALRRSSAAHSTHNTYNTKSENSNRTCSMPYAKLEYKRSS 960
QY 961 NDSLNSVSSNDGYGKRGOMKPSIESYSEDDSEKFCYGOYPADLAHKIHSANEMDNDGE 1020
Db 961 NDSLNSVSSNDGYGKRGOMKPSIESYSEDDSEKFCYGOYPADLAHKIHSANEMDNDGE 1020
QY 1021 LDTPIYSLKYSEBQLNSGRQSPSQNERWARPKHIIIEDEIKQSEQRQSRNQSTTYPVTE 1080
Db 1021 LDTPIYSLKYSEBQLNSGRQSPSQNERWARPKHIIIEDEIKQSEQRQSRNQSTTYPVTE 1080
QY 1081 STDDKHLKFPQHFQOQECVSPYRSRANGSETNRVSGNHGINONVQSLOEDEDYDDKP 1140
Db 1081 STDDKHLKFPQHFQOQECVSPYRSRANGSETNRVSGNHGINONVQSLOEDEDYDDKP 1140
QY 1141 TNSERYSEBEQHEEBERTPNYSIKYNEEKRVDPIDYSLKYATDIPSSQKQSFSSKS 1200
Db 1141 TNSERYSEBEQHEEBERTPNYSIKYNEEKRVDPIDYSLKYATDIPSSQKQSFSSKS 1200
```

1201 SSGSSKTEHSSSENTSTPSSNAKRONQHPSAASRSQPOKAATCKVSSINQETIQ 1260  
 1201 SSGSSKTEHSSSENTSTPSSNAKRONQHPSAASRSQPOKAATCKVSSINQETIQ 1260  
 1261 TYCVEDTPICSRCSLSSLSAASDEIGCNOTTQADSAANTLOAEIKGIGKIGTSAEDPV 1320  
 1261 TYCVEDTPICSRCSLSSLSAASDEIGCNOTTQADSAANTLOAEIKGIGTSAEDPV 1320  
 1321 SEVPAVSQHPRTKSRLOGSSLSSESAPHKAVEPSSGAKSPKSGAQTPKSPPHYVOET 1380  
 1321 SEVPAVSQHPRTKSRLOGSSLSSESAPHKAVEPSSGAKSPKSGAQTPKSPPHYVOET 1380  
 1381 PLMFSTRCTSVSSLSDFERSSTASSVQSEPCSGMWGIIISPSDLDPSPGQTMPPERSKTPP 1440  
 1381 PLMFSTRCTSVSSLSDFERSSTASSVQSEPCSGMWGIIISPSDLDPSPGQTMPPERSKTPP 1440  
 1441 PPPQTAQKREVPKXNKAAPTAKRESGPKQAANAAVQVQLPDADTLLHPATESTPDGF 1500  
 1441 PPPQTAQKREVPKXNKAAPTAKRESGPKQAANAAVQVQLPDADTLLHPATESTPDGF 1500  
 1501 SCSSLSALSDEPTIOKDELRTMPVOENDNGNETESEOPKESNENOEKAEKTIIDSE 1560  
 1501 SCSSLSALSDEPTIOKDELRTMPVOENDNGNETESEOPKESNENOEKAEKTIIDSE 1560  
 1561 KDLLDDSDDDDIIELEECIIISAMPTKSSRKGKPAQATASKLPPPVARKPSQLPVYKLLPS 1620  
 1561 KDLLDDSDDDDIIELEECIIISAMPTKSSRKGKPAQATASKLPPPVARKPSQLPVYKLLPS 1620  
 1621 QNRLOPQKHVSFTPODDMPRYVCEGTPIINSTATSLSDLIIESPPNELAAGEVGRGAQ 1680  
 1621 QNRLOPQKHVSFTPODDMPRYVCEGTPIINSTATSLSDLIIESPPNELAAGEVGRGAQ 1680  
 1681 SGPEFKRTIPIEGRSTDEAOGKTSVTIPELDNKAEEGDIIAECINSAMPKKGSHK 1740  
 1681 SGPEFKRTIPIEGRSTDEAOGKTSVTIPELDNKAEEGDIIAECINSAMPKKGSHK 1740  
 1741 FRVKKIMQVQOASASSAPNKNQDLGKKKPTSPVKIPONTYRTVRKNAKSNLN 1800  
 1741 FRVKKIMQVQOASASSAPNKNQDLGKKKPTSPVKIPONTYRTVRKNAKSNLN 1800  
 1801 ABRVSDNKKSKONLKNNSKDFNDKLNEDRVGSGFAFSPHYTPIEGTPVCFERN 1860  
 1801 ABRVSDNKKSKONLKNNSKDFNDKLNEDRVGSGFAFSPHYTPIEGTPVCFERN 1860  
 1861 SLSSLDFFDDDDVLSREKAELEKAKENKESAKVTSHTELSNQOASANKTQAIKQPINR 1920  
 1861 SLSSLDFFDDDDVLSREKAELEKAKENKESAKVTSHTELSNQOASANKTQAIKQPINR 1920  
 1921 GQPKILOQSTFPOSSKDIIPRGAATDEKLNFAIENTPVCFSHNSLSLSIDQENN 1980  
 1921 GQPKILOQSTFPOSSKDIIPRGAATDEKLNFAIENTPVCFSHNSLSLSIDQENN 1980  
 1981 NKENEPIKETEPDQSGPSKPOASGYAPKSPHVEDTTPVCFERNSSLSLSIDSEDDLLQ 2040  
 1981 NKENEPIKETEPDQSGPSKPOASGYAPKSPHVEDTTPVCFERNSSLSLSIDSEDDLLQ 2040  
 2041 ECISAMPKKKPSRLKGNEXHSPRNMGIIIGEDTLDLKDIOQPDSEHGLSPDSENF 2100  
 2041 ECISAMPKKKPSRLKGNEXHSPRNMGIIIGEDTLDLKDIOQPDSEHGLSPDSENF 2100  
 2101 WKAIQEGANSIYSSLHQAAAACLRQASDDSDSILSLKSGISLSPHLPDQEKFT 2160  
 2101 WKAIQEGANSIYSSLHQAAAACLRQASDDSDSILSLKSGISLSPHLPDQEKFT 2160  
 2161 SNKGRPRILKPGKSTLETKKIESKIGKGGKVKYKSLITGVRNSNBSISGMQKQLOAN 2220  
 2161 SNKGRPRILKPGKSTLETKKIESKIGKGGKVKYKSLITGVRNSNBSISGMQKQLOAN 2220  
 2221 MFSISGRMTIHPGVRNSSSTSPVSKGPKLPTPASKSPSEGTATTSRGAKPSYKS 2280  
 2221 MFSISGRMTIHPGVRNSSSTSPVSKGPKLPTPASKSPSEGTATTSRGAKPSYKS 2280  
 2281 ELSPVARQTSQICGSGKAPSRGSRDSTPSRPAQOPLSRPIQSPGRNSISPGRNGISPPN 2340

RESULT 5

US-08-681-219-32  
 ; Sequence 32, Application US/08681219  
 ; Publication No. US20020058607A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Takaaki Sato and Junn Yanagisawa  
 ; TITLE OF INVENTION: COMPOUNDS THAT INHIBIT THE INTERACTION BETWEEN  
 ; TITLE OF INVENTION: SIGNAL-TRANSDUCING PROTEINS AND THE GLGF  
 ; TITLE OF INVENTION: (PDZ/DHR) DOMAIN AND USES THEREOF  
 ; NUMBER OF SEQUENCES: 35  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Cooper & Dunham LLP  
 ; STREET: 1185 Avenue of the Americas  
 ; CITY: New York  
 ; STATE: New York  
 ; COUNTRY: U.S.A.  
 ; ZIP: 10036  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/681,219  
 ; FILING DATE: 22-JUL-1996  
 ; CLASSIFICATION: 435  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: White, John P  
 ; REGISTRATION NUMBER: 28,678  
 ; REFERENCE/DOCKET NUMBER: 0575/48962/JPW/JKM  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (212) 278-0400  
 ; TELEFAX: (212) 391-0525

2281 ELSPVARQTSQICGSGKAPSRGSRDSTPSRPAQOPLSRPIQSPGRNSISPGRNGISPPN 2340  
 2341 KLSQLPRTSPSTASTKSSGSGKMSYTSPPGMSQONLTQOTGLSKNAGSIIPRESASGK 2400  
 2341 KLSQLPRTSPSTASTKSSGSGKMSYTSPPGMSQONLTQOTGLSKNAGSIIPRESASGK 2400  
 2401 LNMNNGANGANKVELSRMSSTKSGSESDRBERPVLVRQSTFIKEAPSTLRRKLEESA 2460  
 2401 LNMNNGANGANKVELSRMSSTKSGSESDRBERPVLVRQSTFIKEAPSTLRRKLEESA 2460  
 2461 SPESLSPSRPASPTRSQAQTPVLSPLDMSLSTHSSVQAGGWRKLPNLSPTIENDG 2520  
 2461 SPESLSPSRPASPTRSQAQTPVLSPLDMSLSTHSSVQAGGWRKLPNLSPTIENDG 2520  
 2521 RPAKRHDIAARSHESPSRLPINRSGTWKKEHSHSSSLPRVSTWRTGSSSSTLSASSSS 2580  
 2521 RPAKRHDIAARSHESPSRLPINRSGTWKKEHSHSSSLPRVSTWRTGSSSSTLSASSSS 2580  
 2581 SEKAKSEDEKHVNSISGTSQSKENQVSAKGTWRKIKENEFPSTNSTQTVSSGATNGAES 2640  
 2581 SEKAKSEDEKHVNSISGTSQSKENQVSAKGTWRKIKENEFPSTNSTQTVSSGATNGAES 2640  
 2641 KTLIYOMAPAVSKTEDVWVRIEDCPINNPRSGSPGTGNTPPVIDSVSEKANPNIKDKON 2700  
 2641 KTLIYOMAPAVSKTEDVWVRIEDCPINNPRSGSPGTGNTPPVIDSVSEKANPNIKDKON 2700  
 2701 QAKQNVGNSVPMRTVGLNRLTSFIQVADAPDOKGTEIKPGQNNPVVSETNESPIVERT 2760  
 2701 QAKQNVGNSVPMRTVGLNRLTSFIQVADAPDOKGTEIKPGQNNPVVSETNESPIVERT 2760  
 2761 PFSSSSSSKSHSPSGTVAARVTPFNPNPRKSSADSTARSQIPTPVNNNTKKRDSKT 2820  
 2761 PFSSSSSSKSHSPSGTVAARVTPFNPNPRKSSADSTARSQIPTPVNNNTKKRDSKT 2820  
 2821 DSTESSGTOQPKHSGSYLWTSV 2843  
 2821 DSTESSGTOQPKHSGSYLWTSV 2843

; INFORMATION FOR SEQ ID NO: 32:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 2843 amino acids  
 ; TYPE: amino acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: peptide  
 US-08-681-219-32

Query Match 99.5%; Score 14506; DB 8; Length 2843;  
 Best Local Similarity 99.5%; Pred. No. 0;  
 Matches 2830; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

QY	1	MAAASYDQLLKQV	60
DB	1	MAAASYDQLLKQV	60
QY	61	ASSGQIDLLERL	120
DB	61	ASSGQIDLLERL	120
QY	121	RGFVNGSRESTGY	180
DB	121	RGFVNGSRESTGY	180
QY	181	QTDLTRQLEYEAR	240
DB	181	QTDLTRQLEYEAR	240
QY	241	EAERSSQNKHETG	300
DB	241	EAERSSQNKHETG	300
QY	301	PRRLTSLHGTKE	360
DB	301	PRRLTSLHGTKE	360
QY	361	HGNDKDSVLLGN	420
DB	361	HGNDKDSVLLGN	420
QY	421	WEQEAHEPGMDQ	480
DB	421	WEQEAHEPGMDQ	480
QY	481	VDCEMYGLTNDH	540
DB	481	VDCEMYGLTNDH	540
QY	541	QQVIAVLRLNWR	600
DB	541	QQVIAVLRLNWR	600
QY	601	ENKADICAVDGA	660
DB	601	ENKADICAVDGA	660
QY	661	CLQTLQHLKSH	720
DB	661	CLQTLQHLKSH	720
QY	721	GSAALRLNMAN	780
DB	721	GSAALRLNMAN	780
QY	781	PKASHRSKQKH	840
DB	781	PKASHRSKQKH	840
QY	841	LDSSRSSEKOR	900
DB	841	LDSSRSSEKOR	900
QY	901	QEDRSSGTTTEL	960
DB	901	QEDRSSGTTTEL	960
QY	961	NDLSNSVSSNDG	1020
DB	961	NDLSNSVSSNDG	1020
QY	1021	LDTPINYSKYS	1080
DB	1021	LDTPINYSKYS	1080
QY	1081	STDDKHLKFOH	1140
DB	1081	STDDKHLKFOH	1140
QY	1141	TNYSERYSEEB	1200
DB	1141	TNYSERYSEEB	1200
QY	1201	SSGQSKTEHMS	1260
DB	1201	SSGQSKTEHMS	1260
QY	1261	TYCVEDTPICF	1320
DB	1261	TYCVEDTPICF	1320
QY	1321	SEVPAVSQHP	1380
DB	1321	SEVPAVSQHP	1380
QY	1381	PLMFSRCTSV	1440
DB	1381	PLMFSRCTSV	1440
QY	1441	PPPTAQTKRE	1500
DB	1441	PPPTAQTKRE	1500
QY	1501	SCSSLSLAL	1560
DB	1501	SCSSLSLAL	1560
QY	1561	KDLLDDSDDD	1620
DB	1561	KDLLDDSDDD	1620
QY	1621	QNRLOPKHVS	1680
DB	1621	QNRLOPKHVS	1680
QY	1681	SGEPEKDDT	1740
DB	1681	SGEPEKDDT	1740
QY	1741	FRVKIMDVQV	1800
DB	1741	FRVKIMDVQV	1800
QY	1801	AERFVSNKSK	1860
DB	1801	AERFVSNKSK	1860
QY	1861	SLSLDDDDDD	1920
DB	1861	SLSLDDDDDD	1920
QY	1921	GQPKPILOK	1980
DB	1921	GQPKPILOK	1980
QY	1981	NKNEPIKET	2040
DB	1981	NKNEPIKET	2040

1991 NKENPIKETEPDPSQEPKQASGVAPKSFHVEDTPVCFERNSSLSLSISEDDLLQ 2040  
2041 ECISAMPKKKPKRLKDNKHSRPNVGGILGEDTLDKIQRDPSEHGLSPDSENF 2100  
2041 ECISAMPKKKPKRLKDNKHSRPNVGGILGEDTLDKIQRDPSEHGLSPDSENF 2100  
2101 WKATOEKANSTVSSIHQAAAAACLSROASDSDSILSKSGISLGSFPHLTPOEKEPFT 2160  
2101 WKATOEKANSTVSSIHQAAAAACLSROASDSDSILSKSGISLGSFPHLTPOEKEPFT 2160  
2161 SNKGPRILKPEKSTLETETKTESKGIKGGKVKYKSLITGKVRNSSEISGQMKQPILOAN 2220  
2161 SNKGPRILKPEKSTLETETKTESKGIKGGKVKYKSLITGKVRNSSEISGQMKQPILOAN 2220  
2221 MPSISRGRTMTHIPGVNRSSSTSPVSKKGPPLKTPASKSPSEGTATTSPRGAKPSVKS 2280  
2221 MPSISRGRTMTHIPGVNRSSSTSPVSKKGPPLKTPASKSPSEGTATTSPRGAKPSVKS 2280  
2281 ELSPVARTSOIGGSSKAPSGSRDSTPPSPAQPLSRPIQSPGRNSISPGRNGISFPN 2340  
2281 ELSPVARTSOIGGSSKAPSGSRDSTPPSPAQPLSRPIQSPGRNSISPGRNGISFPN 2340  
2341 KLSQLPRTSSPTASTKSGSGKMKYTSFPGRMQOQNLTKQTGLSKVASSIPRSESASG 2400  
2341 KLSQLPRTSSPTASTKSGSGKMKYTSFPGRMQOQNLTKQTGLSKVASSIPRSESASG 2400  
2401 LNMWNGANKVVELSRMSTKSGSDSDSERPVLVROSTFIKEAPSPFLRRKLBESA 2460  
2401 LNMWNGANKVVELSRMSTKSGSDSDSERPVLVROSTFIKEAPSPFLRRKLBESA 2460  
2461 SFESLSPSSRRPASPTRSQAQTPVLSPLPDMSLTHSSVQAGGWRKLPPLNLSPTIEYNDG 2520  
2461 SFESLSPSSRRPASPTRSQAQTPVLSPLPDMSLTHSSVQAGGWRKLPPLNLSPTIEYNDG 2520  
2521 RPAKHDIARSHESPSRLPTNRSCTWKREHSKSSSLPVSTWERTGSSSSILSASSES 2580  
2521 RPAKHDIARSHESPSRLPTNRSCTWKREHSKSSSLPVSTWERTGSSSSILSASSES 2580  
2581 SEKAKSEDEKHVNSTSGTKQKENVASAKGTWRKIKENEFSPNTSTSTVSSGATNGAES 2640  
2581 SEKAKSEDEKHVNSTSGTKQKENVASAKGTWRKIKENEFSPNTSTSTVSSGATNGAES 2640  
2641 KTLIIQMAPAVSKTDDVWVREDCEINNPRSGRPTGNTPTVIDSVSEKANPNIKSKDN 2700  
2641 KTLIIQMAPAVSKTDDVWVREDCEINNPRSGRPTGNTPTVIDSVSEKANPNIKSKDN 2700  
2701 QAKQVNGSVPMRTVGLNRLTSTIQVADPQKTEIKPGONNPVPVSEINNEPIVERT 2760  
2701 QAKQVNGSVPMRTVGLNRLTSTIQVADPQKTEIKPGONNPVPVSEINNEPIVERT 2760  
2761 PPSSSSSKHSGSPSGTVAARVTPFNPNPRKSSADTSARPSQIPTPVNNNTKRDST 2820  
2761 PPSSSSSKHSGSPSGTVAARVTPFNPNPRKSSADTSARPSQIPTPVNNNTKRDST 2820  
2821 DSTESSGTSQPKRHSGLVTSV 2843  
2821 DSTESSGTSQPKRHSGLVTSV 2843

## RESULT 6

US-09-230-111C-30  
; Sequence 30, Application US/09230111C  
; Publication No. US20030203414A1  
; GENERAL INFORMATION:  
; APPLICANT: Sato, Taka-Aki  
; TITLE OF INVENTION: COMPOUNDS THAT INHIBIT INTERACTION BETWEEN  
; TITLE OF INVENTION: SIGNAL-TRANSDUCING PROTEINS AND THE GLCF (PDZ/DHR)  
; TITLE OF INVENTION: DOMAIN AND USES THEREOF  
; FILE REFERENCE: 48962-A-PCT-US  
; CURRENT APPLICATION NUMBER: US/09/230,111C  
; CURRENT FILING DATE: 1999-05-17

NUMBER OF SEQ ID NOS: 33  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 30  
; LENGTH: 2843  
; TYPE: PRT  
; ORGANISM: human  
US-09-230-111C-30  
Query Match 99.5%; Score 14506; DB 11; Length 2843;  
Best Local Similarity 99.5%; Pred. No. 0;  
Matches 2830; Conservative 5; Mismatches 8; Indels 0; Gaps 0;  
QY 1 MAAASVDQLLKQVEALKMENSNLROELEDNSNHLTKLETEASNMYKVLKQSGSIEDEAM 60  
Db 1 MAAASVDQLLKQVEALKMENSNLROELEDNSNHLTKLETEASNMYKVLKQSGSIEDEAM 60  
QY 61 ASSGQIDLLERLKEINLNDSSNPPGVKLRSMKLSRYSVSGREGSVSRSGSCSPVPMGSFPR 120  
Db 61 ASSGQIDLLERLKEINLNDSSNPPGVKLRSMKLSRYSVSGREGSVSRSGSCSPVPMGSFPR 120  
QY 121 RGFVNGSRESTGYLELEKERSILLADLDKEKEKDWYVAQNLTKRIDSPLPTENFSL 180  
Db 121 RGFVNGSRESTGYLELEKERSILLADLDKEKEKDWYVAQNLTKRIDSPLPTENFSL 180  
QY 181 QDLTRQLEYEARQIRVAMEBOLGTCQDMEKRAQRRIARIQOIEKDIIRIROLLOSQAT 240  
Db 181 QDLTRQLEYEARQIRVAMEBOLGTCQDMEKRAQRRIARIQOIEKDIIRIROLLOSQAT 240  
QY 241 EADRSQNKHETGSHDAERQNEGQGVGEINMATSNGQSGSTTRMDHETASVLSSSSTHSA 300  
Db 241 EADRSQNKHETGSHDAERQNEGQGVGEINMATSNGQSGSTTRMDHETASVLSSSSTHSA 300  
QY 301 PRRLTSHLGTQVEMVYSLLSMLGTHDKDDMSRTLLAMSSSDSCISMRSGCCLPLLIQL 360  
Db 301 PRRLTSHLGTQVEMVYSLLSMLGTHDKDDMSRTLLAMSSSDSCISMRSGCCLPLLIQL 360  
QY 361 HGMNDKSVLLGNRSGKEARASAAHNIHSDPPDKRGRREIRVHLLEQIRAYCETC 420  
Db 361 HGMNDKSVLLGNRSGKEARASAAHNIHSDPPDKRGRREIRVHLLEQIRAYCETC 420  
QY 421 WEMQEAHEPGMDQKPNMPAPVEHQICPAVCVLMKLSFDEEHRHANNEGLQAIABELQ 480  
Db 421 WEMQEAHEPGMDQKPNMPAPVEHQICPAVCVLMKLSFDEEHRHANNEGLQAIABELQ 480  
QY 481 VDCEMYGLTNDHYSITLRRYAGNALNTLPGDVANKATLCSMGCMRALVAQLKSSSEDL 540  
Db 481 VDCEMYGLTNDHYSITLRRYAGNALNTLPGDVANKATLCSMGCMRALVAQLKSSSEDL 540  
QY 541 QQVIASVLRNLSWRADVNSKKTLEEVGSVKALMECALEVKESTLKSVALNLSAHC 600  
Db 541 QQVIASVLRNLSWRADVNSKKTLEEVGSVKALMECALEVKESTLKSVALNLSAHC 600  
QY 601 ENKADICAVDGAFLVGLTYRSQNTLAIIESGGGILRNVSLLIATNEDHRCILRENN 660  
Db 601 ENKADICAVDGAFLVGLTYRSQNTLAIIESGGGILRNVSLLIATNEDHRCILRENN 660  
QY 661 CLQTLLOHLKSHSLTIVSNACGTLNLSARNPKDQALWDMGAVSMLKNIHSGKHMIAM 720  
Db 661 CLQTLLOHLKSHSLTIVSNACGTLNLSARNPKDQALWDMGAVSMLKNIHSGKHMIAM 720  
QY 721 GSAAALNLMANRPAPKYDANIMSPGSSLSLHVRRKQKALEAELDAQHSETFDNINLS 780  
Db 721 GSAAALNLMANRPAPKYDANIMSPGSSLSLHVRRKQKALEAELDAQHSETFDNINLS 780  
QY 781 PKASHRSKQRHKQSLYDGYFDNTRHDDNRSDFNTGNMTVLPYLNTTVLPSSSSSRGS 840  
Db 781 PKASHRSKQRHKQSLYDGYFDNTRHDDNRSDFNTGNMTVLPYLNTTVLPSSSSSRGS 840  
QY 841 LDDSRSEKRSLENERGIGLVNHPATENPGTSKRGLOISTTAAQIAKVMEEVSAIHTS 900  
Db 841 LDDSRSEKRSLENERGIGLVNHPATENPGTSKRGLOISTTAAQIAKVMEEVSAIHTS 900  
QY 901 QEDRSSGSTTEHLHCVTDERNALRSSAAHTSNITNTFKSENSNRTCSMPYAKLEYKRSS 960



Db 901 QEDRSSGTTLEHCVTDERNALRSSAAHTSNFTKSENSNRTCSMPYAKLEYKRSS 960  
Qy 961 NDSLNSVSSNDGYKRGOMKPSIESYEDDESKEFCSYGOYPADLAHIAHNSANWMDNDGE 1020  
Db 961 NDSLNSVSSNDGYKRGOMKPSIESYEDDESKEFCSYGOYPADLAHIAHNSANWMDNDGE 1020  
Qy 1021 LDTINTSLKYSDDLNSGRQSPQNRWARPKHIIIEDEIKQSEQRNRQSTTPVYTE 1080  
Db 1021 LDTINTSLKYSDDLNSGRQSPQNRWARPKHIIIEDEIKQSEQRNRQSTTPVYTE 1080  
Qy 1081 STDDKHLKFPQHFQOCVSPYRSGANGSETNRVGSNHGNGVNSOSLCOEDDYEDDKP 1140  
Db 1081 STDDKHLKFPQHFQOCVSPYRSGANGSETNRVGSNHGNGVNSOSLCOEDDYEDDKP 1140  
Qy 1141 TNSERYSEEEQHEBERPTNYSIKYNEEKRHVDQPIDYSLKYATDIPSSQKQSFSSKS 1200  
Db 1141 TNSERYSEEEQHEBERPTNYSIKYNEEKRHVDQPIDYSLKYATDIPSSQKQSFSSKS 1200  
Qy 1201 SSGOSSKTEHMSSESTSTPSSNAKQONOLHPSSAQSRGQPKAATCKVSSINQETIQ 1260  
Db 1201 SSGOSSKTEHMSSESTSTPSSNAKQONOLHPSSAQSRGQPKAATCKVSSINQETIQ 1260  
Qy 1261 TYCVEDTPICFSRCSLSSLSAEDIEGNCQTTQEADSAANTLOIAEIKGIGTRSAEDPV 1320  
Db 1261 TYCVEDTPICFSRCSLSSLSAEDIEGNCQTTQEADSAANTLOIAEIKGIGTRSAEDPV 1320  
Qy 1321 SEYPAUSCHRTKSSRLOGSLSESARHKAIVPPSGAKSPKSGAQTPKSPPEHYVQET 1380  
Db 1321 SEYPAUSCHRTKSSRLOGSLSESARHKAIVPPSGAKSPKSGAQTPKSPPEHYVQET 1380  
Qy 1381 PLMFSRCTSVSSLDSPESRGIASVQSEPCSGMVSGIISPSDLPDGPOTMPPSRKTPP 1440  
Db 1381 PLMFSRCTSVSSLDSPESRGIASVQSEPCSGMVSGIISPSDLPDGPOTMPPSRKTPP 1440  
Qy 1441 PPQTAQTKREVPRNKAPTAKEKESGPKQAANVAORVQVLDPADTLHFATESTPDGF 1500  
Db 1441 PPQTAQTKREVPRNKAPTAKEKESGPKQAANVAORVQVLDPADTLHFATESTPDGF 1500  
Qy 1501 SCSSLSALSLEDPFIQVELRIMPPVQENDNGNETSEOPKESNENOEKAETIDSE 1560  
Db 1501 SCSSLSALSLEDPFIQVELRIMPPVQENDNGNETSEOPKESNENOEKAETIDSE 1560  
Qy 1561 KOLLDDSDDDDIIEILBECIISAMPTKSSRKGKPAQTASKLPPPVARKPSQLPVYKLLPS 1620  
Db 1561 KOLLDDSDDDDIIEILBECIISAMPTKSSRKGKPAQTASKLPPPVARKPSQLPVYKLLPS 1620  
Qy 1621 QNRLOPKQHVSTPGDDMPRVYCVETGPIINFSTATSLDITTESPNELAAAGEVGRGAQ 1680  
Db 1621 QNRLOPKQHVSTPGDDMPRVYCVETGPIINFSTATSLDITTESPNELAAAGEVGRGAQ 1680  
Qy 1681 SGFEKXEDTIPTREGRSTDEAQQGKTSSVTIPELDDNKAEEGDIABEINSAMPKQKSHKP 1740  
Db 1681 SGFEKXEDTIPTREGRSTDEAQQGKTSSVTIPELDDNKAEEGDIABEINSAMPKQKSHKP 1740  
Qy 1741 FRVKKINDVQOASASSAPNKNQDGGKKKTPSPVKPIPQNTYRTRVRKNADSKNNLN 1800  
Db 1741 FRVKKINDVQOASASSAPNKNQDGGKKKTPSPVKPIPQNTYRTRVRKNADSKNNLN 1800  
Qy 1801 AERVFSNDKSKONLKNNSKDFNDKLPNNEDRVGSAFDPGPHVTPTEGTPYCFSRND 1860  
Db 1801 AERVFSNDKSKONLKNNSKDFNDKLPNNEDRVGSAFDPGPHVTPTEGTPYCFSRND 1860  
Qy 1861 SLSSLDFFDDDDVLSREKAEKAKENKSEAKVTSHTELTNQOQSAKNTQIAKQPINR 1920  
Db 1861 SLSSLDFFDDDDVLSREKAEKAKENKSEAKVTSHTELTNQOQSAKNTQIAKQPINR 1920  
Qy 1921 GQPKPILQKQSTFPQSSKDIIPDRGAATDEKLNFAIENTPVCFSHNSLSLSDIDQENN 1980  
Db 1921 GQPKPILQKQSTFPQSSKDIIPDRGAATDEKLNFAIENTPVCFSHNSLSLSDIDQENN 1980  
Qy 1981 NKNEPIKETEPPDDSGGSEPSKPAQSGYAPKSFHVEDTTPVCFSRNSSLSLSDIDEDDLLQ 2040

Db 1981 NKNEPIKETEPPDDSGGSEPSKPAQSGYAPKSFHVEDTTPVCFSRNSSLSLSDIDEDDLLQ 2040  
Qy 2041 ECISAMPKKKPSRLKGDNEKHSPRNMGGLGEDLTLDLKDIOBPDSHGLSPDSENF 2100  
Db 2041 ECISAMPKKKPSRLKGDNEKHSPRNMGGLGEDLTLDLKDIOBPDSHGLSPDSENF 2100  
Qy 2101 WKAIQEGANSIVSSLHQAAAAACLSQASSSDSLSLKGISGLSPPHLTPDQEEKPFT 2160  
Db 2101 WKAIQEGANSIVSSLHQAAAAACLSQASSSDSLSLKGISGLSPPHLTPDQEEKPFT 2160  
Qy 2161 SNKGPRILKPEKSTLETTKIESKGIKGGKVKYKSLITGKVRNSHSEISQMKQPLQAN 2220  
Db 2161 SNKGPRILKPEKSTLETTKIESKGIKGGKVKYKSLITGKVRNSHSEISQMKQPLQAN 2220  
Qy 2221 MPSISRGRTMIHIPGVNRSSSTSPVSKKGPPLKTPASKSPSEGTATTSPRGAKPSVKS 2280  
Db 2221 MPSISRGRTMIHIPGVNRSSSTSPVSKKGPPLKTPASKSPSEGTATTSPRGAKPSVKS 2280  
Qy 2281 ELSIPVAROTSOIGSSSKAPSRSGSRDSTPSRPAQOPLSRPIQSPGRNSISPRNGISPPN 2340  
Db 2281 ELSIPVAROTSOIGSSSKAPSRSGSRDSTPSRPAQOPLSRPIQSPGRNSISPRNGISPPN 2340  
Qy 2341 KLSQLPRTSSPSTASTKSSGSKMSYTPGRQMSQOQLTKQTGLSKNASIPIRSESASKG 2400  
Db 2341 KLSQLPRTSSPSTASTKSSGSKMSYTPGRQMSQOQLTKQTGLSKNASIPIRSESASKG 2400  
Qy 2401 LNQMNGANGANKVELGRMSTKSSGSESDRSEPRVLRQSTFTIKEAPSPTLRKLBEA 2460  
Db 2401 LNQMNGANGANKVELGRMSTKSSGSESDRSEPRVLRQSTFTIKEAPSPTLRKLBEA 2460  
Qy 2461 SPESLSPSSRPASTRQOATPVLSPLPDMSLSTHSSVOAGGWRKLPNLSPTIEYNDG 2520  
Db 2461 SPESLSPSSRPASTRQOATPVLSPLPDMSLSTHSSVOAGGWRKLPNLSPTIEYNDG 2520  
Qy 2521 RPAKHDIARSHSESPRLPINRSGTWKREHSHSSLPVSTWRTGSSSSILSASSES 2580  
Db 2521 RPAKHDIARSHSESPRLPINRSGTWKREHSHSSLPVSTWRTGSSSSILSASSES 2580  
Qy 2581 SEKAKSDEKHVNSISGTQKSKENOVSAKGTWRKIKENEFSTNSTQVSSGATNGAES 2640  
Db 2581 SEKAKSDEKHVNSISGTQKSKENOVSAKGTWRKIKENEFSTNSTQVSSGATNGAES 2640  
Qy 2641 KTLIQVAPAVSKTEDVWVRIEDCPINNPRSGSPTCNTPPVIDSVSEKANPNIKSKDN 2700  
Db 2641 KTLIQVAPAVSKTEDVWVRIEDCPINNPRSGSPTCNTPPVIDSVSEKANPNIKSKDN 2700  
Qy 2701 QAKQVNGSVPMRTVGLNRLTSFIOVDAPDQKGTETIKPGQNNPVVSTNESPIVERT 2760  
Db 2701 QAKQVNGSVPMRTVGLNRLTSFIOVDAPDQKGTETIKPGQNNPVVSTNESPIVERT 2760  
Qy 2761 PFSSSSSKSHSSPSGTVAAARVTPFNYPNPSPRKSSADSTARPQIPTPVNNNTKGRDST 2820  
Db 2761 PFSSSSSKSHSSPSGTVAAARVTPFNYPNPSPRKSSADSTARPQIPTPVNNNTKGRDST 2820  
Qy 2821 DSTESSGTQSPKXHSGLYVTSV 2843  
Db 2821 DSTESSGTQSPKXHSGLYVTSV 2843

## RESULT 7

US-10-092-138-30  
; Sequence 30, Application US/10092138  
; Publication No. US20030170723A1  
; GENERAL INFORMATION:  
; APPLICANT: Sato, Taka-Aki  
; TITLE OF INVENTION: METHOD OF PREPARING A PROTEIN ARRAY BASED ON  
; TITLE OF INVENTION: BIOCHEMICAL PROTEIN-PROTEIN INTERACTION  
; FILE REFERENCE: 65823/JPW/PT  
; CURRENT APPLICATION NUMBER: US/10/092,138  
; NUMBER OF SEQ ID NOS: 33  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 30



QY	961	NDLSNVSSNDGYGKRGOMKPSIESYSEDESKFCSYQYPADLAHKIHSANHMDNDGE	1020
DB	961	NDLSNVSSNDGYGKRGOMKPSIESYSEDESKFCSYQYPADLAHKIHSANHMDNDGE	1020
QY	1021	LDTPIYNSLYKDEQLNSGRQSPQNERWARPKHIEDEIKOSEQOSQNSQSTTPYVTE	1080
DB	1021	LDTPIYNSLYKDEQLNSGRQSPQNERWARPKHIEDEIKOSEQOSQNSQSTTPYVTE	1080
QY	1081	STDHKLKQPHFGQOECVSPYRSRGANGSETRVSGNHNQNVQSQSCQEDDYEDDKP	1140
DB	1081	STDHKLKQPHFGQOECVSPYRSRGANGSETRVSGNHNQNVQSQSCQEDDYEDDKP	1140
QY	1141	TNYSERYSEEBEHEERPTNYSIKYNEBKRVHDQIDYSLKYANDIPSSQKQSPFSKS	1200
DB	1141	TNYSERYSEEBEHEERPTNYSIKYNEBKRVHDQIDYSLKYANDIPSSQKQSPFSKS	1200
QY	1201	SSQCSKTEHMSSTNTSTPSSNAKRONQLPSSAQSSGPOKAAATCKVSINQETIQ	1260
DB	1201	SSQCSKTEHMSSTNTSTPSSNAKRONQLPSSAQSSGPOKAAATCKVSINQETIQ	1260
QY	1261	TYCVETPTICFSCSSLSLSSAEDEIGCQNTTQEADSAANTLQIAEIKIGITRSEADPV	1320
DB	1261	TYCVETPTICFSCSSLSLSSAEDEIGCQNTTQEADSAANTLQIAEIKIGITRSEADPV	1320
QY	1321	SEYPAYSOHPRTKSSRLQSSLSSESARHKAVEFPSSGAKSPSKGAQTPKSPPEHYVQET	1380
DB	1321	SEYPAYSOHPRTKSSRLQSSLSSESARHKAVEFPSSGAKSPSKGAQTPKSPPEHYVQET	1380
QY	1381	PLMFSRCTSVSSLDSPESRSIASSVQSEPCSGMVSGIISPSDLPSPGQTMPPSRKTPP	1440
DB	1381	PLMFSRCTSVSSLDSPESRSIASSVQSEPCSGMVSGIISPSDLPSPGQTMPPSRKTPP	1440
QY	1441	PPQTAQTKREVPKNKAPTAEKRESGPKAAVNAAVQVQLPADTLLHFAESTPDGF	1500
DB	1441	PPQTAQTKREVPKNKAPTAEKRESGPKAAVNAAVQVQLPADTLLHFAESTPDGF	1500
QY	1501	SCSSLSALSLDSPFFIQKVELRIMPVQNDNGNETESEQPKESNENQKEAEKIDSE	1560
DB	1501	SCSSLSALSLDSPFFIQKVELRIMPVQNDNGNETESEQPKESNENQKEAEKIDSE	1560
QY	1561	KDLSDSDDDDDIIELEECIISAMPTKSSRKGAQPAQTAASKLPPPVARKPSQLPVYKLLPS	1620
DB	1561	KDLSDSDDDDDIIELEECIISAMPTKSSRKGAQPAQTAASKLPPPVARKPSQLPVYKLLPS	1620
QY	1621	QNLQPKHVSFTPGDDMPRVYCVETGPTINFSTATSLSDLTIESPPNELAAGVGVRGAQ	1680
DB	1621	QNLQPKHVSFTPGDDMPRVYCVETGPTINFSTATSLSDLTIESPPNELAAGVGVRGAQ	1680
QY	1681	SGEFKEDTTIPTEGRSTDEAQQGKTSSVTIPELDDNKAEGDILABCINSAMPKGSHP	1740
DB	1681	SGEFKEDTTIPTEGRSTDEAQQGKTSSVTIPELDDNKAEGDILABCINSAMPKGSHP	1740
QY	1741	FRVKKINDQVQOASASSAPNKNQDCKKKKPTSPVKPIPQNTTEYRTRVRKNADSKNLN	1800
DB	1741	FRVKKINDQVQOASASSAPNKNQDCKKKKPTSPVKPIPQNTTEYRTRVRKNADSKNLN	1800
QY	1801	AERFSDNCKSKONLKNNSKDFNDKLPNNEDRVGSGFAPDSPHHYTPIEGTFCFSRND	1860
DB	1801	AERFSDNCKSKONLKNNSKDFNDKLPNNEDRVGSGFAPDSPHHYTPIEGTFCFSRND	1860
QY	1861	SLSLDFFDDDDVLSREKAEIRKAKENKESEAKVTSHTELTSQOOSANKTAQAKPINR	1920
DB	1861	SLSLDFFDDDDVLSREKAEIRKAKENKESEAKVTSHTELTSQOOSANKTAQAKPINR	1920
QY	1921	GQPKIILQKSTFPQSSKDIPIDRGAATDEKLQNFATENTPVCFSHNSLSLSLSDIDQENN	1980
DB	1921	GQPKIILQKSTFPQSSKDIPIDRGAATDEKLQNFATENTPVCFSHNSLSLSLSDIDQENN	1980
QY	1981	NKNEPIKETETPPDSQGEPSKPAQSGYAPKSFHVETPPVCFSRNNSLSLSISIDEDLLQ	2040
DB	1981	NKNEPIKETETPPDSQGEPSKPAQSGYAPKSFHVETPPVCFSRNNSLSLSISIDEDLLQ	2040
QY	2041	ECISSAMFKKKKPRLLKGDNEKHSRPNWGGILGDLTLDKIQRPDSEHGLSPDSENF	2100

QY	1	MAAASYDOLLKQVEALKMNSNLQLEEDNHNH.TKLETEASNKVVLKQLOGSIEDEAM	60
DB	1	MAAASYDOLLKQVEALKMNSNLQLEEDNHNH.TKLETEASNKVVLKQLOGSIEDEAM	60
QY	61	ASSGOIDLLERLKEINLIDSSNFPVGVKRSKMSLASYSRSGESVSRSGECPVPMGSPFR	120
DB	61	ASSGOIDLLERLKEINLIDSSNFPVGVKRSKMSLASYSRSGESVSRSGECPVPMGSPFR	120
QY	121	RGFVNGSRSTGYLEELPKERSLLIADLDKEEKDWTYAOLQNLTKRIDSPLTENFSL	180
DB	121	RGFVNGSRSTGYLEELPKERSLLIADLDKEEKDWTYAOLQNLTKRIDSPLTENFSL	180
QY	181	QTDLTRRQLEVEARQIRVAMEEQGLTQDMKRAQRRIARIQOIEKDIRLQQLQSOAT	240
DB	181	QTDLTRRQLEVEARQIRVAMEEQGLTQDMKRAQRRIARIQOIEKDIRLQQLQSOAT	240
QY	241	EAERSQNKHETGSHDARQNEQGVGEINMATSGQGGSTRMDHETASVLSSSSTHSA	300
DB	241	EAERSQNKHETGSHDARQNEQGVGEINMATSGQGGSTRMDHETASVLSSSSTHSA	300
QY	301	PRRLTSHLGTKEVWYSLLSMLGTHDKDDMSRTLLAMSSQDSCTISXQSCCLPLLIQLL	360
DB	301	PRRLTSHLGTKEVWYSLLSMLGTHDKDDMSRTLLAMSSQDSCTISXQSCCLPLLIQLL	360
QY	361	HGNDKXSVLLGNRSGSKARARASAAALHNIHSQPDCKRGRREIRVLLHLLQIRAYCETC	420
DB	361	HGNDKXSVLLGNRSGSKARARASAAALHNIHSQPDCKRGRREIRVLLHLLQIRAYCETC	420
QY	421	NEWQBAHPGMDQXNPAPVPHQICPAVCVLMKLSFDEHRHNMNELGLQIAELLQ	480
DB	421	NEWQBAHPGMDQXNPAPVPHQICPAVCVLMKLSFDEHRHNMNELGLQIAELLQ	480
QY	481	VDCEMYGLTNDHYSITLRYAGMALTNITFGDVANKATLCMKGCMRALVAQLKSEBDL	540
DB	481	VDCEMYGLTNDHYSITLRYAGMALTNITFGDVANKATLCMKGCMRALVAQLKSEBDL	540
QY	541	QOVIASVLRNLSWRADVNSKTLREVGVSKALMECALEVYKKESTLKSVALNLSAHT	600
DB	541	QOVIASVLRNLSWRADVNSKTLREVGVSKALMECALEVYKKESTLKSVALNLSAHT	600
QY	601	ENKADI CAVDGALAFVGLTTRVRSOTNTLAIIESGGIILRVNSSLIATNEDHRLRENN	660
DB	601	ENKADI CAVDGALAFVGLTTRVRSOTNTLAIIESGGIILRVNSSLIATNEDHRLRENN	660
QY	661	CIQTLLOHLKSHSLTIVSNACGTLNLSARNPKDQEALWDMGAVSMLKNLHSHKMIAM	720
DB	661	CIQTLLOHLKSHSLTIVSNACGTLNLSARNPKDQEALWDMGAVSMLKNLHSHKMIAM	720
QY	721	GSAALRNLMNRPKAKYDANIMSGSSLP.SLHVKKQKALEAELDAOHLSETFDNIDLS	780
DB	721	GSAALRNLMNRPKAKYDANIMSGSSLP.SLHVKKQKALEAELDAOHLSETFDNIDLS	780
QY	781	PKASHRSKORHKQSLYGDVYVFTNRRHDDNRSDFNTGNMTVLSPLYNTITVLPSSSSSRGS	840
DB	781	PKASHRSKORHKQSLYGDVYVFTNRRHDDNRSDFNTGNMTVLSPLYNTITVLPSSSSSRGS	840
QY	841	LQSSREKORSERERIGLIGNYHPATENPTGSSKRGLOISITAAQIAKMWEEVSAIHTS	900
DB	841	LQSSREKORSERERIGLIGNYHPATENPTGSSKRGLOISITAAQIAKMWEEVSAIHTS	900
QY	901	QEDRSSGSGTTELHCVTDERNALLRRSAAHTNTYNTFTKSENRTCSMPYAKLYKRS	960
DB	901	QEDRSSGSGTTELHCVTDERNALLRRSAAHTNTYNTFTKSENRTCSMPYAKLYKRS	960

LENGTH: 2843  
TYPE: PRT  
ORGANISM: human  
US-10-092-138-30

Query Match 99.5%; Score 14506; DB 14; Length 2843;  
Best Local Similarity 99.5%; Pred. No. 0;  
Matches 2830; Conservative 5; Mismatches 8; Indels 0; Gaps 0;



QY 1021 LDTPIYSLKYSDOLNSGRSPSONERWAPKHHIIEDEIKQSORQSRNQSTTYPVYTE 1080  
DB 1019 LDTPIYSLKYSDOLNSGRSPSONERWAPKHHIIEDEIKQSORQSRNQSTTYPVYSE 1078  
QY 1081 STDDKHLKQFHFQGOECVSPYRGRGANGSTNRVGNHGNQVNSOSLQCEDDYEDDKP 1140  
DB 1079 NTDDKHLKQFHFQGOECVSPYRGRGANGSTNRVGNHGNQVNSOSLQCEDDYEDDKP 1138  
QY 1141 TNYSERVSEEBQH-EEBERPNYISKYNEEKHVDQPIDYSLKYATDIPSSQKOSFQFSK 1199  
DB 1139 TNYSERVSEEBQH-EEBERPNYISKYNEEKHVDQPIDYSLKYATDIPSSQKOSFQFSK 1198  
QY 1200 SSSGSSCTEHWSSSENSTPSSNAKRONLHPSSAQSRSGOPKKAATCKVSSINQETI 1259  
DB 1199 NSSAQSTKPELSESEENTAVPPNAKQONLRSSAQ-RNGOTQKGTTCVPSINQETI 1257  
QY 1260 QTYCVEDTPICFRCSLSLSSAEDEIGCQNTTQEADSANLQIAEIKGIGTRSAEDP 1319  
DB 1258 QTYCVEDTPICFRCSLSLSSADDEIGCQNTTQEADSANLQIAEIKGIGTRSAEDP 1317  
QY 1320 VSEVPANVCHRTKSRRLQSSLSSESARH-KAYEFPAGAKSPKSGAOTPKSPPEHYVQ 1378  
DB 1318 ATEVPANVQNAKAPSRLOAGSLSESTRHNKAYEFSSGAKSPKSGAOTPKSPPEHYVQ 1377  
QY 1379 ETPLMFSRCTSVSSLDSPESRSIASSVQSEPCSGMVSGIISPSDLPDPSPGOTMPPSRSKT 1438  
DB 1378 ETPLMFSRCTSVSSLDSPESRSIASSVQSEPCSGMVSGIISPSDLPDPSPGOTMPPSRSKT 1437  
QY 1439 PPPPQTAQTRKVPKNAKPAETAKRESGPKQAANAAVQRVQVLPDADTLHPATESTPD 1498  
DB 1438 PPPPQTAQTRKVPKNAKPAETAKRESGPKQAANAAVQRVQVLPDADTLHPATESTPD 1497  
QY 1499 GPCSSSSLSLSDPESRSIASSVQSEPCSGMVSGIISPSDLPDPSPGOTMPPSRSKT 1558  
DB 1498 GPCSSSSLSLSDPESRSIASSVQSEPCSGMVSGIISPSDLPDPSPGOTMPPSRSKT 1556  
QY 1559 SEKDLSDSDDDDDIIEBECIISAMPTKSRKGGKPAQATASKLPPPVARKPSQLPVYKLL 1618  
DB 1557 SEKDLSDSDDDDDIIEBECIISAMPTKSRKGGKPAQATASKLPPPVARKPSQLPVYKLL 1616  
QY 1619 PSQNLQKQKVSFTPGDDMRVYCEVGTPIFNSTATSLSLDTIESPNEELAGEGVGG 1678  
DB 1617 PAQNLQKQKVSFTPGDDMRVYCEVGTPIFNSTATSLSLDTIESPNEELAGEGVGG 1676  
QY 1679 AQSGFEKEDTPTBGRSTDDAQRGKISSIVTPDLDDNKAEEGDILAEICINSAMPKGKSH 1738  
DB 1677 AQSGFEKEDTPTBGRSTDDAQRGKISSIVTPDLDDNKAEEGDILAEICINSAMPKGKSH 1736  
QY 1739 KPFRVKKIMQVQKASASSAPKNQOLGKKKPTSPVKPIPONTEYTRVRKNAQSKN 1798  
DB 1737 KPFRVKKIMQVQKASASSAPKNQOLGKKKPTSPVKPIPONTEYTRVRKNAQSKN 1796  
QY 1799 LNAERVFSDNKSQKONLKNNSKDFNDKLPNNEDRVGRSFAFDSPHHYTPIEGTPYCFPSR 1858  
DB 1797 VNTETTFSDNKSQKPSLQTKAKAFNEKLPNNEDRVGRGTFFALDSPHHYTPIEGTPYCFPSR 1856  
QY 1859 NDSLSLDFDDDDVLSREKAEFLRKAENKSEAKVTSHTELTNSQOSANKTOIAKOPI 1918  
DB 1857 NDSLSLDFDDDDVLSREKAEFLRKAENKSEAKVTSHTELTNSQOSANKTOIAKOPI 1916  
QY 1919 NRGQPKILOKOSTPQSKDIPDRGAATDEKLONFALIENTVPCFNSHNSLSLSDIDQE 1978  
DB 1917 NRGQPKILOKOSTPQSKDIPDRGAATDEKLONFALIENTVPCFNSHNSLSLSDIDQE 1976  
QY 1979 -NNKNEPEIKETEPPDSCPSKPAQAGYAPKPSFHVEDTTPVCFSRNSSLSSLSIDSD 2037  
DB 1977 NNNKNEPEIKETEPPDSCPSKPAQAGYAPKPSFHVEDTTPVCFSRNSSLSSLSIDSD 2036  
QY 2038 LLOECISSAMPKKKPSRLKGNKHSPRNMGGILGEDTLIDLKDIQRPDSEHGLSPDSE 2097  
DB 2037 LLOECISSAMPKKKPSRLKGNKHSPRNMGGILGEDTLIDLKDIQRPDSEHGLSPDSE 2096

QY 2098 NFDWKAIORGANSIYSSLHQ-AAAAACLSROASSSDSILSKSGISLGSPHLPDQEE 2156  
DB 2097 NFDWKAIORGANSIYSSLHQ-AAAAACLSROASSSDSILSKSGISLGSPHLPDQEE 2156  
QY 2157 KPFTSNKGRPILKPKBESKSTLETKIESEKGIKGGKVKYKSLITGKVRNSENISGOMQP 2216  
DB 2157 KPFTSNKGRPILKPKBESKSTLETKIESEKGIKGGKVKYKSLITGKVRNSENISGOMQP 2216  
QY 2217 LOANPESIRGRTHIHPVRNSSSSTSPVSKKGGPLKTPASKSPSEGTATTPREGAKP 2276  
DB 2217 LPNMPESIRGRTHIHPVRNSSSSTSPVSKKGGPLKTPASKSPSEGTATTPREGAKP 2276  
QY 2277 SVKSELSPVARQTSQIGSSKAPSRSGSDSTPSPAQOPLSRPIQSPGRNISPGRNGI 2336  
DB 2277 AGKSELSPITRQTSQISGNGKSSRSGSDSTPSPAQOPLSRPIQSPGRNISPGRNGI 2336  
QY 2337 SPNKLSQLPRTSSPSTASTKSSGGSKMSYTFPGROMSQOONLTQOTGLSKNASSIPRSES 2396  
DB 2337 SPNKLSQLPRTSSPSTASTKSSGGSKMSYTFPGROMSQOONLTQOTGLSKNASSIPRSES 2396  
QY 2397 ASKGLNOMNGNGANKVELSRMSTKSSGSDSERPVLVRQSTFIKEAPSPILRRKL 2456  
DB 2397 ASKGLNOMNGNGANKVELSRMSTKSSGSDSERPVLVRQSTFIKEAPSPILRRKL 2456  
QY 2457 EESASFESLSPSRPASPTRSQATPVLSPSLPDMSLTHSSVQAGWKLPPNLSPTIE 2516  
DB 2457 EESASFESLSPSRPASPTRSQATPVLSPSLPDMSLTHSSVQAGWKLPPNLSPTIE 2516  
QY 2517 YNDGRPAKHDIARSHESPSRLPINESCTWKEHSHKSSSLPRVSTWERTGSSSILSA 2576  
DB 2517 YNDGRPAKHDIARSHESPSRLPINESCTWKEHSHKSSSLPRVSTWERTGSSSILSA 2576  
QY 2577 SSESSEKAKSEDEKHVNSISGTQSKENOVSAKGTWRKIKENEFSTNSTQSTVSSGATN 2636  
DB 2577 SSESSEKAKSEDEKHVNSISGTQSKENOVSAKGTWRKIKENEFSTNSTQSTVSSGATN 2636  
QY 2637 GAESKTLIQMAPAVSKTEDVWVRIEDCPINNPRSCRSPTGNTPPVIDSVSEKANPNKD 2696  
DB 2637 GAESKTLIQMAPAVSKTEDVWVRIEDCPINNPRSCRSPTGNTPPVIDSVSEKANPNKD 2696  
QY 2697 SKDQO--AKONGVNGSPMRTVGLNRLTSTFIQVDAPQKGTGKPGQNNPVPVSETNE 2753  
DB 2697 SKDQO--AKONGVNGSPMRTVGLNRLTSTFIQVDAPQKGTGKPGQNNPVPVSETNE 2753  
QY 2754 SPIVERTPSSSSSKHSPSGVGAARVTFNPNPRKSSADSTARSQIPTPVNNT 2813  
DB 2754 SPIVERTPSSSSSKHSPSGVGAARVTFNPNPRKSSADSTARSQIPTPVNNT 2813  
QY 2814 KKDSDKTDSTESSGTOSEKHSGLVLTSV 2843  
DB 2814 KKDSDKTDSTESSGTOSEKHSGLVLTSV 2843

RESULT 9  
US-09-987-482-2  
; Sequence 2, Application US/09987482  
; Publication No. US20020184656A1  
; GENERAL INFORMATION:  
; APPLICANT: BHANDARI, POONAM  
; APPLICANT: SHASHIDHARA, L.S.  
; TITLE OF INVENTION: IN VIVO ASSAY SYSTEM FOR SCREENING AND VALIDATION OF  
; TITLE OF INVENTION: DRUGS AND OTHER SUBSTANCES  
; FILE REFERENCE: 056859-0134  
; CURRENT APPLICATION NUMBER: US/09/987,482  
; CURRENT FILING DATE: 2002-03-21  
; NUMBER OF SEQ ID NOS: 3  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 2  
; LENGTH: 912  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-987-482-2

Query Match 32.4%; Score 4719; DB 9; Length 912;  
Best Local Similarity 99.6%; Pred. No. 8.2e-259;  
Matches 908; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 959 SSNDLSNSVNDGYGRGQKQSIYSYSDDESKSCYQYPADLAHKIHSANHMDND 1018  
DB 1 SSNDLSNSVSSDGYGRGQKQSIYSYSDDESKSCYQYPADLAHKIHSANHMDND 60

QY 1019 GELDTPIYSLKXSDQLNSGRSPQNERWARPKHIIIEDEIKQSEQRNRQSTTYPYV 1078  
DB 61 GELDTPIYSLKXSDQLNSGRSPQNERWARPKHIIIEDEIKQSEQRNRQSTTYPYV 120

QY 1079 TESTDDKHLKXFOHFGQOECVSPYRSGANGSTNRVGNHGINQVNSLOCEDDYEDD 1138  
DB 121 TESTDDKHLKXFOHFGQOECVSPYRSGANGSTNRVGNHGINQVNSLOCEDDYEDD 180

QY 1139 KPTNYSERYSEEEHHEEERPTNYSIKYNEEKHVVQPIDYSLKYATDIPSSQKQSFSPS 1198  
DB 181 KPTNYSERYSEEEHHEEERPTNYSIKYNEEKHVVQPIDYSLKYATDIPSSQKQSFSPS 240

QY 1199 KSSGGQSKTEHMSSESSENTSTPSSNAKRONQVLPSSAQSRSQPOKAATCKVSSINQET 1258  
DB 241 KSSGGQSKTEHMSSESSENTSTPSSNAKRONQVLPSSAQSRSQPOKAATCKVSSINQET 300

QY 1259 IQTYCVEDTPICRSCSSLSLSAEDIEICNQTTOEADSANTLQIAEIKGIGTRASD 1318  
DB 301 IQTYCVEDTPICRSCSSLSLSAEDIEICNQTTOEADSANTLQIAEIKGIGTRASD 360

QY 1319 PVSEVPVAVSQHPTKSRQLSGSLSESARHKAVEFPFGSAKSPSKSGAOTPKSPPEHYQ 1378  
DB 361 PVSEVPVAVSQHPTKSRQLSGSLSESARHKAVEFPFGSAKSPSKSGAOTPKSPPEHYQ 420

QY 1379 ETPLMFSCRTSVSLDSFERSSTASSVQSEPCGMVSGIISPSDLPSQGMTPERSKT 1438  
DB 421 ETPLMFSCRTSVSLDSFERSSTASSVQSEPCGMVSGIISPSDLPSQGMTPERSKT 480

QY 1439 PPPPQTAQTKREVPKNKAFTAEKRESGPKQAANAQVQVLPDADTLHFATESTPD 1498  
DB 481 PPPPQTAQTKREVPKNKAFTAEKRESGPKQAANAQVQVLPDADTLHFATESTPD 540

QY 1499 GFSCSSLSALSDEPPIQKDVLRIMPVQENDNGNETESEQPKESNENQEKAKTID 1558  
DB 541 GFSCSSLSALSDEPPIQKDVLRIMPVQENDNGNETESEQPKESNENQEKAKTID 600

QY 1559 SEKDLDSDSDDDIEILEECIIISAMPTKSRKKGKPAQATSKLPPPVARKPSQLPYKLL 1618  
DB 601 SEKDLDSDSDDDIEILEECIIISAMPTKSRKKGKPAQATSKLPPPVARKPSQLPYKLL 660

QY 1619 PSQNRLOPQKHVFTPGDDMPRYVCVEGTPINFSTATSLDITSPPNELAAEGVRGG 1678  
DB 661 PSQNRLOPQKHVFTPGDDMPRYVCVEGTPINFSTATSLDITSPPNELAAEGVRGG 720

QY 1679 AQSGEPFKRTIITEGRSTDEAGGKTSSVTIPELDNKAEEGDIIAECINSAMPKGS 1738  
DB 721 AQSGEPFKRTIITEGRSTDEAGGKTSSVTIPELDNKAEEGDIIAECINSAMPKGS 780

QY 1739 KPRVKKIMQVOQAASASSAPKNQLDGKKKPTSPVKPIPONTYRTRVRKNADSKN 1798  
DB 781 KPRVKKIMQVOQAASASSAPKNQLDGKKKPTSPVKPIPONTYRTRVRKNADSKN 840

QY 1799 LNAERFVFNKDSKKNLKNNSKDFNDKLPNNEDVRGSGFADSPHHYTPTEGTPYCFSR 1858  
DB 841 LNAERFVFNKDSKKNLKNNSKDFNDKLPNNEDVRGSGFADSPHHYTPTEGTPYCFSR 900

QY 1859 NDSLSSLDFFDD 1870  
DB 901 NDSLSSLDFFDD 912

RESULT 10

US-09-987-482-3

; Sequence 3, Application US/09987482

; Publication No. US20020184656A1

GENERAL INFORMATION:  
; APPLICANT: BHANDARI, POONAM  
; APPLICANT: SHASHIDHARA, L.S.  
; TITLE OF INVENTION: IN VIVO ASSAY SYSTEM FOR SCREENING AND VALIDATION OF  
; TITLE OF INVENTION: DRUGS AND OTHER SUBSTANCES  
; FILE REFERENCE: 056859-0134  
; CURRENT APPLICATION NUMBER: US/09/987,482  
; CURRENT FILING DATE: 2002-03-21  
; NUMBER OF SEQ ID NOS: 3  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 3  
; LENGTH: 767  
; TYPE: PRN  
; ORGANISM: Homo sapiens  
US-09-987-482-3

Query Match 26.7%; Score 3889; DB 9; Length 767;  
Best Local Similarity 99.9%; Pred. No. 7.7e-212;  
Matches 766; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAASYSQQLLKQVEALKMENSNLQLEEDNSNHLTKLETSANMKEVLKOLQGSIEDEAM 60  
DB 1 MAASYSQQLLKQVEALKMENSNLQLEEDNSNHLTKLETSANMKEVLKOLQGSIEDEAM 60

QY 61 ASSGQIDLLERLKLNDSSNFPQVKLRKMSLSYSGRSYSSRSRSGECSVPVMSGFP 120  
DB 61 ASSGQIDLLERLKLNDSSNFPQVKLRKMSLSYSGRSYSSRSRSGECSVPVMSGFP 120

QY 121 RGFVNGSRESTGYLEELKERSLLADLDKEEKQWYLAQLQNLTKRIDSLPTENFSL 180  
DB 121 RGFVNGSRESTGYLEELKERSLLADLDKEEKQWYLAQLQNLTKRIDSLPTENFSL 180

QY 181 QTDLTRQLEVEARQIVAMEEQGTGQDMEKRAQRIARIQIEKDIILRIQLQSQAT 240  
DB 181 QTDLTRQLEVEARQIVAMEEQGTGQDMEKRAQRIARIQIEKDIILRIQLQSQAT 240

QY 241 EAERSSQNKHETGSHDAERQNEGQVGEINMATSNGCGSTTRMDHETASVLSSTHSA 300  
DB 241 EAERSSQNKHETGSHDAERQNEGQVGEINMATSNGCGSTTRMDHETASVLSSTHSA 300

QY 301 PRRLTSHLGTKEVWYVYLLSMLGTHDKDMSRTLLANSSQDSCISMRQSGCLPLLIQL 360  
DB 301 PRRLTSHLGTKEVWYVYLLSMLGTHDKDMSRTLLANSSQDSCISMRQSGCLPLLIQL 360

QY 361 HGNDKDSVLLGNSRSGKEARASAAALHNIHSOPDDKRGREIRVLHLEQIRAYCETC 420  
DB 361 HGNDKDSVLLGNSRSGKEARASAAALHNIHSOPDDKRGREIRVLHLEQIRAYCETC 420

QY 421 WEWOEAHEPGMDQKMPAPVEHQICPAVCVLMKLSFDEEHRHAMNELGGLQAIALLQ 480  
DB 421 WEWOEAHEPGMDQKMPAPVEHQICPAVCVLMKLSFDEEHRHAMNELGGLQAIALLQ 480

QY 481 VDCSMYGLTNDHYSITLRRYAGVATLNLTFQDVANKATLCSMKGCMRALVAQLKSEEDL 540  
DB 481 VDCSMYGLTNDHYSITLRRYAGVATLNLTFQDVANKATLCSMKGCMRALVAQLKSEEDL 540

QY 541 QQVIASVLRNLSWPAADVNSKKTUREVGSVKALMECALEVKESTLKSLSALNLSAHCT 600  
DB 541 QQVIASVLRNLSWPAADVNSKKTUREVGSVKALMECALEVKESTLKSLSALNLSAHCT 600

QY 601 ENKADIICAVGALAFVLTITYSQNTLAIISGGGILNRVSSLIATNEDHQIILRENN 660  
DB 601 ENKADIICAVGALAFVLTITYSQNTLAIISGGGILNRVSSLIATNEDHQIILRENN 660

QY 661 CLQTLQHLKSHSLTIYSNACGTLNLSARNPKDQEAALWDMGAVSMLKNLIHSHKHMAM 720  
DB 661 CLQTLQHLKSHSLTIYSNACGTLNLSARNPKDQEAALWDMGAVSMLKNLIHSHKHMAM 720

QY 721 GSAALRNLANRPAKYKDANIMSPGSSLSLHVRKOKALEAELDAQ 767  
DB 721 GSAALRNLANRPAKYKDANIMSPGSSLSLHVRKOKALEAELDAQ 767

```

RESULT 11
US-10-267-502-373
; Sequence 373, Application US/10267502
; Publication No. US20040071700A1
; GENERAL INFORMATION:
; APPLICANT: Kim, Jaeseob
; APPLICANT: Galant, Ron
; TITLE OF INVENTION: Obesity Linked Genes
; FILE REFERENCE: LSD-07416
; CURRENT APPLICATION NUMBER: US/10/267,502
; CURRENT FILING DATE: 2003-01-27
; NUMBER OF SEQ ID NOS: 439
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 373
; LENGTH: 2274
; TYPE: PRT
; ORGANISM: Mus musculus
; US-10-267-502-373

Query Match      24.1%; Score 3508.5; DB 12; Length 2274;
Best Local Similarity 34.4%; Pred No. 1.4e-189;
Matches 1005; Conservative 356; Mismatches 789; Indels 773; Gaps 98

QY      4  ASYDQLLKQVEALKMKNENLNLSQELNEDSNHLTKLETASNMKEVLKQLOGSIEDEA--MA 61
Db      6  ASVEQLVRQVEALKAEKNTHLRQELRDNSHLSKLETETSGMKEVLKHLQKLEQEARVLV 65

QY      62  SSGQIDLLERKELNLDSSNFPGVKLRSKLSLSYSGREGSVSSRSGECSPYPMGSPFPR 121
Db      66  SSGQTEVLEQLKALQTDLSYYLNKLFAP----ALGPEP---AARTPEGSPV-HGSGPSK 117

QY      122 -GFVNGSRESTGYLEELEKERSLLIADLKEEKEKWYQAQLQMLTKRIDSPLTENFSL 180
Db      118  DSFGELSRATIRLLEELQDERCFLLSEIEKEEKLWYYSQQLGSKRLDELPHVDVTSM 177

QY      181  QTDLTRQLREYARQIRVAMEEQIGTQDNKEKAQRRIARIQIEKDIL-----RIQQ--- 233
Db      178  QMDLIRQLLEFEAQHRSLEMEERFCTSDENVQRAIRASELEQIKELLEAQRDYQOTEP 237

QY      234  --LLOSQATEAERSQNHETGSHDAERQNEGQGVGINMATSGNGQGTTRMDHETASV 291
Db      238  QALLAVKPVAVEEQEABVTHPEDGTPQ-----PGN----- 269

QY      292  LSSSTHSAPRLTSHLGTQKVMYVYLLSMGLTHDKDMSRTLLAMSSQDSCISNROSG 351
Db      270  -----SKVEVVFLLSLMLATRDQEDTARTLLAMSSSPESCVAMRRSG 311

QY      352  CLPLLIQLLHGNKDQSV--LLNGNSRGSKEARAPASAAHLNIHSDPDKKGRREIRVLH 408
Db      312  CLPLQLQLHGTGTEAGSVGRAGIPGAPGAKARMRANAHLNIHVSDFDGLARKEMRVLI 371

QY      409  LLEQIRAYCETCWEOQEAHEFGMDQKNPVPAPVEHQICPAVCVLMKLSFDEEHRHAMNE 468
Db      372  VLEQIRAYCETCWDMLQARDSGTE---TPVPIEPQICQATCAVMKLSFDEEYRRAMNE 426

QY      469  LGLGQAIAEILLOVDCEMYGLTNDHYSITLRYAGMALTNITFGDVANKATILCSMKGCWRA 528
Db      427  LGLGQAVAEILQVDYEMHKMTRDPLNLALRYAGMTLNLITFGDVANKATLCARRGCMEA 486

QY      529  LVAQLKSSEDLQQVIAQVNLNLSWRADVNSKKTLEVGSVKALMECALEVKESTLKSIV 588
Db      487  IVAQLGSSEELHQVVSILNLSWRADINSKKVLEVGSMTALMECVLRASKESTLKSIV 546

QY      589  LSAWNLNSAHTENKADICAVDGAIAFLVGTLYRSTNTLAIIESGGILNRVSSLIAT 648
Db      547  LSAWNLNSAHTENKAAICQVDGALGFLVSTLYRCCGNSLAVIESGGGILNRVSSLIAT 606

QY      649  NEDHRQILRENNCLQTLLQHLKSHSLTIVTNSACCTLNLSARNPKDQEAALWDMGAVMLK 708
Db      607  REDIYQVLRDHNCLQTLLQHLTSHSLTIVTNSACCTLNLSARSPROEQLLWDLGAVMLR 666

QY      709  NLIHSKXKVIAMGSAALNLMANRPAYKDANI-KSPGSSLSPLHVRKQKALEAELDAQ 767

```







RESULT 13  
 US-10-267-502-368  
 ; Sequence 368, Application US/10267502  
 ; Publication No. US20040071700A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Kim, Jaeseob  
 ; APPLICANT: Galant, Ron  
 ; TITLE OF INVENTION: Obesity Linked Genes  
 ; FILE REFERENCE: LSD-07416  
 ; CURRENT APPLICATION NUMBER: US/10/267,502  
 ; CURRENT FILING DATE: 2003-01-27  
 ; NUMBER OF SEQ ID NOS: 439  
 ; SOFTWARE: Patent version 3.2  
 ; SEQ ID NO 368  
 ; LENGTH: 2417  
 ; TYPE: PRT  
 ; ORGANISM: Drosophila melanogaster  
 US-10-267-502-368

Query Match 12.1%; Score 1763; DB 12; Length 2417;  
 Best Local Similarity 25.3%; Pred. No. 1.6e-90;  
 Matches 720; Conservative 326; Mismatches 836; Indels 964; Gaps 101;

QY	243	ERSSONKHETGSHDAERON-EGQGVGEINMATSGGQSTTR-----MDHETASVLSSS	295
DB	129	ELREMRHRSIDRNFERSQAOQQQDLPPRNGGGSPASAGRPSRSEKPSYTLRSFLDGD	188
QY	296	STHSAPR-----RLTSH-----LGTKVMVYLLSMLGTHDKDMSRTLLAM	337
DB	189	APAPAPRLPKGAAMTTGDFERYTSSAVEATLGSKEVCYVILLSMLGSDNDPLEMAKKFLEL	248
QY	338	SSSQDCISWRQSCPLPILQLLHGNDKDSVLLNGSEKSEARASAAALHNIHSQPD	397
DB	249	SGNAQSCATLRSSCMPLLVQMHPND-----QEVKRCAGQALHNVVSHPOE	298
QY	398	KRGREIRVLHLLQIRAYCETCWEQAEHPGM--DQDNMPAPVVEHOICPAVCVLMK	455
DB	299	KAGREAKVRLDQIVDYCSFLKTLQSGGEALADSDRHPL-----AAISSLMK	349
QY	456	LSPDEERHAWNELGGLOAELLQVDCENYV-LTNDHYSITLRRVAGHALTNLTFGDVA	514
DB	350	VSPDEERHAWNELGGLOAELLQVDCENYV-LTNDHYSITLRRVAGHALTNLTFGDVA	514
QY	515	NKATLCSMKGOMRALVAQLSESDLOQVIAVLRNLNWRADVNSKTLREVGSVKALME	574
DB	410	NKALLCCQKQFMALVAQLSDAPDQLLOVTAIVLRNLNWRADSNMKAVALNEIGTVTALAL	469
QY	575	CALEVKESTLKVLSALWNLASHCTENKADI CAVGALAFVLTLYTSRQNTLALIES	634
DB	470	AAMNRSNTLKATLSALWNLASHCTENKADI CAVGALAFVLTLYTSRQNTLALIES	634
QY	635	GGGILRVNLSIATNEDHROILRENQCLOTLLOLKHSHSLTIIVSNACGTLNLSARNPKD	694
DB	530	AGGILRVNLSHIAVCEPYRQILQHNCALAILLOQLKSESITVVSNSCGTLNLSARNSD	589
QY	695	QEALWDNGAVSMLKNLHSHKXMIAMGSAALRNLMANRPARYK-----DANIMSPG-SSL	749
DB	590	QKFLWDNGAVSMLRSLHSHKXMIAMGSAALRNLMANRPARYK-----DANIMSPG-SSL	749
QY	750	PSLHVKKOKALELDAQHLSETFDNIDNLSPKASHRSKQRHKQSLYGYVDFTNRHDDN	809
DB	650	PTLEARKAKAQQLGEBRHTAETCDNLD-----	677
QY	810	RSDNFNTGNMTVLSPLYNTITVLPSSSSSGSLDSRSSEKDRSLERBERGICLGNYPATEN	869
DB	678	-----TGG-----KLDKERASSSSRRHP-----	695
QY	870	PGTSSKRGLOISTTAQIAKVMEEVSAIHTSQDRSSGSGTTELHCYVTDENALRRSSAAH	929
DB	696	-----APLPLR-----SAMLTKSSSRDSVVSYSKASDCAVD--HLIRASASD	734

QY	930	THSNTYNTFTKGENSNRTCSMPYAKLEYKRSSNDLSNVSSNDGYGKRGQMKPSIESYSED	989
DB	735	AHR-----KVRPKITDFDLE	749
QY	990	DESKFCSYGOYPADLAHKLHSANEMDNDGELDTPIYNSLYSKYSEQLNSGRQSPQSNRHW	1049
DB	750	ME-----QDTEATEQPIDYSVKYS-----	769
QY	1050	ARPKHIIIEIKQSEQRQNRQSTTYPVYTESTDDKHLKQFHPFGQOCVSPYRSRGANG	1109
DB	770	-----ENAKTSTY-----	778
QY	1110	SETNRVGNHGINQNVOSLQEDDYEDDKPTNYSERYSEEQHEBERPTNYSIKNEE	1169
DB	779	-----QETDL--DQPTDFSLRYAE-----NQIESDLDISGPAGGQ	811
QY	1170	KRHVDQPIDYSLKYATDIP--SSQKQSFSSKSSGQSKTEH-----MSSSENSTSPSS	1223
DB	812	KSTITPP-----AETVPEKSEGEIILLDDSVKCYQETEDTPYVISNAASVTDLRVAA	864
QY	1224	NKRONQLHPS--SAQSRSGOPO--KAATCKVSSINOETIQTVCVEDTPICFGRCSLS	1278
DB	865	KADAEAYKPEVREVTSKGAPKPLKSLQCGSGSYTPEKPINYCEGTGPGYFGRYDLS	924
QY	1279	SLSSAEDIEIG--CNOTTQEADSANTLQIAEIKGIGTSAEDPVEVPAVSQHPRTKSSRL	1337
DB	925	SL-----DESGKANQIIVGTD-----ADIKPLEKEQE-----	954
QY	1338	QGSLSSSASARKHAFVPPSGAKSPKSGAOTPKSPPEHYVOETPLMFSTRCTSVSSLSDFE	1397
DB	955	-----SQPAEQVLTKPPTQANS-----ALETPLMFSRSSMDSLVHDP	992
QY	1398	SESTA-----SSVQSEPCSGMVSGIISPSLDPSDPTGTPMPSRSKTPPPPTQATXREV	1452
DB	993	DVDVANCDDKSSVSD--FSRLASGVISPEIPDPTQSMPOS-----	1033
QY	1453	PKNKAPTAEKESGPKQAAVNAVQVLPDADTLHFFATESTPDGFCSSLSALS--	1511
DB	1034	PRNSVAGSGQNVDSPPVWPASLOPLRSVFE--DLSSFNVEHTPAQFSTATSLNSLSIV	1092
QY	1512	DEPIQKXDELRLMPPVOENDNGNE-----TESEOPKESNENQEK--AKET	1556
DB	1093	DDE-----KAPASVAEEDNEDELLANCMQKRPTEAVKSTVNVSEVDVAEET	1143
QY	1557	IDS-----EKD-----LLDD--SDDDI--	1572
DB	1144	IRSYCTEDTPALLSKVPSNTNLSVISMSSTDPKDATAGQMAHQSLDSDVSSNASDCGG	1203
QY	1573	---EILECIIISAMPTKSSRKKGKPAOTASKLPPPVARK--PSQLPVYKLLPSONRLOPQK	1628
DB	1204	ASGHLLQOCIRDMG-----KKPLGEATSDPTAMLRGNGNELPGY--LPS-----	1245
QY	1629	HVSTPTGDMPRVYCVGCTPIINFSTATSLDLTIES-----PPNELAAGEGVGGAQSGEF	1684
DB	1246	-----ADEKXK--FLVEDSPCNFVSVGLNLTGVSSLVGPVQLKETE--PSSADQNPEN	1297
QY	1685	EKRPTIPEGSTDEAQGGKTSVVTIPELDDNKABEGDILAE-----	1727
DB	1298	KAKPGQEQVRRPWHQDDSLSSIDSDEDTNLLSQAIAGCNRPKSNLGFSSNGKRSS	1357
QY	1728	-----INSAMEKGSKHKPFVKIMQVQOASASSAPKNQO-----LDG--	1767
DB	1358	SLSSSQPIAANAATSSALNSAMTVRSQQOSESYSVSDSDSDNDNQSKSLFELCILKGY	1417
QY	1768	KKKKP-----TSPVKPIQNTYEYRT-----RVRKXADSKNLNA	1801
DB	1418	KTKEPGARAQQMQBPVIGSSVSQNSPLKQFSLPVQLPSSGGQVKQRHHHHHHRRER	1477
QY	1802	ERVFSDNKDSK--KQNLKNN--SKDFNDKLPNNEDVRGSEAFDPSPHYTYETPTCYFSR	1858
DB	1478	ER-----ERKDEKLLQECINTGISKKIN--AVPKNV--LATSAAALEPCH--PMAAT--	1523
QY	1859	NDLSLSLDFDDDDVDLSREKAEALFKAKENKESAKVTSTHTLTSTNQOSANKTOAIKQPI	1918



Db 1524 -TSASALSTAAPDV-----EOKAHATSNPQ---QSSHPSSHILNPI 1563  
Qy 1919 -----NRQPKPILOKOSTPQSSKOIPDRGAATDEKLNPALEN 1558  
Db 1564 DAIATVTDIVRSPAAPNGNGNSQNGLET-ATGSKOLDSDRSDSNOSFINETWVRL 1622  
Qy 1959 -----TPVCF-----SHNSLSLSLSDIDQENN 1980  
Db 1623 DSALNETICISASEKHDPDLMLKSVBLTMEFTVTSAEQLRSSSHNHS 1677  
Qy 1981 KNEPIKETEPSPQSGPKQAGVAPKSFHVEDTPVCFNRNSLSLSID---SEDD 2037  
Db 1678 NSSNNTWNESTCPNDVSPFSVQTA-----PV-----LASLSLDEDATEAR 1718  
Qy 2038 LLOECISSAMPKKKPSKLGDNK-----HSPRNGGILGEDLTLDLKDIORPSEHGLS 2093  
Db 1719 SLHELIEITPNEQOPESEGETDLVNGHADSVSGSGGLNFQJ-----GGQVQAGVRL 1775  
Qy 2094 PDSENFWDKAIQEGANSIVSSJHQAACLSRQASDSD-----SLSL-----KS 2140  
Db 1776 FORLLFNCTASIMTNTMIAFEAPALAEENLLQPAATDDDTTMTFSLNSLDLONIRPPS 1835  
Qy 2141 GI-SLGSFP--HLTPDQEKPTSNKGRILPKGEKSTLETKTESBKIGKGGKVKYS 2197  
Db 1836 GMSLNSCYQDHSQSSLSURQAMPS--KSPRFARKMPPANLVARALGHLAG----- 1884  
Qy 2198 LITGVRNSISGOMKOP-----LOANPSTISGRWTIHIPVRNSSSTSPVSKKGP 2252  
Db 1885 -SAESVNSCNLLDNIPKPSLMDLDSMISVDSTQSEVADGEQDCSMATTISVSNYETA 1943  
Qy 2253 L-----KTPASKPSPEGCTATTTPRGAKPSVKSEL----- 2282  
Db 1944 ACDDQMTVLOSCEDEDEATMNDYSAAESTPKHGSTPSNRRSLTPKDKERLTAKDRPKT 2003  
Qy 2283 -----SPVARTSQI-----GGSKAPSRSGSRDSTPSRPAQQLSPRI----- 2321  
Db 2004 YTIATSCWEAPEANETLQIEIVEAAVPVATPSFRANGRRGSAERYKTLIECPALIQ 2063  
Qy 2322 -----KTPASKPSPEGCTATTTPRGAKPSVKSEL-----QSP--GRNSISPR 2333  
Db 2064 POPDDCPSEQLSSIRAMQOFTFITDINIGHSQETCESTDHPEDAGESPECQNSSETSC 2123  
Qy 2334 NG-----TSPPNKLSQL-----PTSSPSTASTKSSGSGKMSYTSPCGROMSQOQLTKQT 2382  
Db 2124 DGOEPDQPPPPSIVDLRTSVVKFTLEBPATAVKLVGRKKPAYVSPYMSQSRN----- 2178  
Qy 2383 GLSKNASSIPRESASAKGLNQMNGNGANKKVELSRMSTKS-----SGSESRSRSPV- 2436  
Db 2179 -----SNNAPSKKTLPTIAKRSVLPGSGVRLPAKKKTP 2216  
Qy 2437 -----LVROSTTIKEAPSTLRRKLEESASFESLSPSRPASPTRSQAOTPVLSPL 2488  
Db 2217 PPEPARLERQGTFTKD-----EPTNSNVQVPVVE-TK 2249  
Qy 2489 PMSLSHSSVQAGWRKLPPLNLSPTIEYNDGRPAKRHDIAHSHSESPSRLPINRSGTWK 2548  
Db 2250 PAQTSPTHRA-----SKLPTKKGTA---SGSPSK-----AGSKRPIPLAFA---R 2289  
Qy 2549 REHSKHSSLPVSTWRRTGSSSLSASSESEKAKSEDEKHNVSISGTSKQENQVSA 2608  
Db 2290 RMTPOKANTSLAAGKSHAASRVSVSSTTPPSRNSNLNGSSAAAAAKINHAQS 2349  
Qy 2609 K-GTRWKIKENEFSTNS---TSQTVSSGATNGAESKTLIYQMAPAVSKTEDVWRIED 2663  
Db 2350 RIANIWRVDEAKTKOSSNLNLTQTKYSSNMLNANGTKFTLRL-----SSTFD----- 2397  
Qy 2664 CPINNPRSGRSPGTGTPPVVSDVSEK 2689  
Db 2398 -----NTPSTAGGVKSK 2409

RESULT 14

US-09-915-307-5  
; Sequence 5, Application US/09915307  
; Publication No. US20020015943A1  
; GENERAL INFORMATION:  
; APPLICANT: Bienz, Marian  
; TITLE OF INVENTION: Assays, methods and means relating to the modulation of  
; TITLE OF INVENTION: levels of nuclear beta-catenin  
; FILE REFERENCE: SMW/FP5864806  
; CURRENT APPLICATION NUMBER: US/09/915,307  
; CURRENT FILING DATE: 2002-11-22  
; PRIOR APPLICATION NUMBER: US 60/221,892  
; PRIOR FILING DATE: 2000-07-31  
; NUMBER OF SEQ ID NOS: 12  
; SOFTWARE: Patent in Ver. 2.1  
; SEQ ID NO 5  
; LENGTH: 325  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-915-307-5  
  
Query Match 11.3%; Score 1640; DB 12; Length 325;  
Best Local Similarity 99.4%; Pred. No. 8.3e-85;  
Matches 323; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
  
Qy 1263 CVEDTPTCSRCSLSLSLSAEDIEICNQTOTQADSANTLQIAEIKGIGTRSAEDPVSE 1322  
Db 1 CVEDTPTCSRCSLSLSAEDIEICNQTOTQADSANTLQIAEIKGIGTRSAEDPVSE 60  
Qy 1323 VPAVSQHPRTKSRLOGSSLSSESARHKAVEFPFGAKSPSKGAQTPKSPPEHYVQETPL 1382  
Db 61 VPAVSQHPRTKSRLOGSSLSSESARHKAVEFPFGAKSPSKGAQTPKSPPEHYVQETPL 120  
Qy 1383 MFRCTSVSLDSFERSSTASSVQSPFCGMSVGIISPSDLPSGQTMPPSPSKTTPPPP 1442  
Db 121 MFRCTSVSLDSFERSSTASSVQSPFCGMSVGIISPSDLPSGQTMPPSPSKTTPPPP 180  
Qy 1443 PQAQTKREVPKNAKPTAEKRESGPKQAANAAVQVQLPDAITLLHFATESTPDGFSC 1502  
Db 181 PQAQTKREVPKNAKPTAEKRESGPKQAANAAVQVQLPDAITLLHFATESTPDGFSC 240  
Qy 1503 SSSLSALSDEPTIKDVELRIMPVQENDNGNETSEQPKESNENQKEAKETIDSEKD 1562  
Db 241 SSSLSALSDEPTIKDVELRIMPVQENDNGNETSEQPKESNENQKEAKETIDSEKD 300  
Qy 1563 LLDSDDDDDIEILEECIIISAMPTKS 1587  
Db 301 LLDSDDDDDIEILEECIIISAMPTKS 325

RESULT 15  
US-10-029-386-33090  
; Sequence 33090, Application US/10029386  
; Publication No. US20030194704A1  
; GENERAL INFORMATION:  
; APPLICANT: Penn, Sharon G.  
; APPLICANT: Hanzel, David K.  
; APPLICANT: Hanzel, David K.  
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR CH  
; TITLE OF INVENTION: EXPRESSION ANALYSIS TWO  
; FILE REFERENCE: ACOMICA-X-2  
; CURRENT APPLICATION NUMBER: US/10/029,386  
; CURRENT FILING DATE: 2001-12-20  
; NUMBER OF SEQ ID NOS: 34288  
; SOFTWARE: Annonax Sequence Listing Engine vers. 1.1  
; SEQ ID NO 33090  
; LENGTH: 1633  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: MAP TO AC027307.3  
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.85  
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 0.84  
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 4.8



Search completed: August 25, 2004, 17:52:37  
Job time : 216.5 secs

**This Page Blank (uspto)**

Core version 5.1.6  
1993 - 2004 Compugen Ltd.

ing sw model

, 17:08:42 ; Search time 37.5 Seconds

(without alignments)

3913.937 Million cell updates/sec

2

EAALMEN.....ESSGTQSPKRGSGSYLTVSV 2843

Text 0.5

625971 residues

Chosen parameters: 389414

000

0%

100%

45 summaries

AA:\*

data/2/iaa/5A COMB.pdp:\*

data/2/iaa/5B COMB.pdp:\*

data/2/iaa/6A COMB.pdp:\*

data/2/iaa/6B COMB.pdp:\*

data/2/iaa/PCTUS COMB.pdp:\*

data/2/iaa/backfiles1.pdp:\*

Of results predicted by chance to have a  
ual to the score of the result being printed,  
is of the total score distribution.

SUMMARIES

ID	Description
1	US-07-741-940-2
1	Sequence 2, Appli
1	US-08-289-548A-2
1	Sequence 2, Appli
1	US-08-452-654-2
1	Sequence 2, Appli
2	US-08-370-235A-2
1	Sequence 2, Appli
1	US-08-449-731-2
1	Sequence 2, Appli
1	US-08-452-655B-2
1	Sequence 7, Appli
1	US-08-452-655B-7
1	Sequence 2, Appli
3	US-08-450-582-7
1	Sequence 7, Appli
2	US-08-821-355A-7
1	Sequence 7, Appli
2	US-09-003-687A-7
1	Sequence 7, Appli
2	US-09-136-605-7
1	Sequence 7, Appli
1	US-07-741-940-7
1	Sequence 7, Appli
1	US-08-289-548A-7
1	Sequence 7, Appli
1	US-08-452-654-7
1	Sequence 7, Appli
4	US-08-449-731-7
1	Sequence 7, Appli
2	US-08-630-822A-100
2	Sequence 100, App
2	US-09-005-069-100
1	Sequence 100, App
4	US-09-171-156A-49
1	Sequence 49, Appl
4	US-09-004-730A-49
1	Sequence 49, Appl
4	US-08-981-799A-49
1	Sequence 49, Appl
4	US-09-134-001C-4463
1	Sequence 4463, Ap
4	US-09-134-001C-3159
1	Sequence 3159, Ap
4	US-09-134-001C-5080
1	Sequence 5080, Ap
1	US-08-353-700-1
1	Sequence 1, Appli
5	PCT-US95-16216-1
1	Sequence 1, Appli
4	US-08-714-741-32
1	Sequence 32, Appli

28	371.5	2.5	1581	3	US-09-110-517-2	Sequence 2, Appli
29	371.5	2.5	3256	4	US-09-919-172-98	Sequence 98, Appli
30	371.5	2.5	3256	4	US-09-976-594-22	Sequence 22, Appli
31	371.5	2.5	3969	3	US-08-061-376-5	Sequence 5, Appli
32	371	2.5	2468	4	US-09-976-594-726	Sequence 726, App
33	355	2.4	2662	4	US-09-595-684B-31	Sequence 31, Appli
34	336	2.3	2954	4	US-09-150-867-1	Sequence 1, Appli
35	334.5	2.3	2185	4	US-09-854-856-36	Sequence 36, Appli
36	334.5	2.3	2245	4	US-09-854-856-4	Sequence 4, Appli
37	334.5	2.3	2332	4	US-09-854-856-34	Sequence 34, Appli
38	334.5	2.3	2382	4	US-09-854-856-2	Sequence 2, Appli
39	326.5	2.2	2482	1	US-08-328-254-6	Sequence 6, Appli
40	321	2.2	2157	4	US-09-854-856-52	Sequence 52, Appli
41	321	2.2	2217	4	US-09-854-856-20	Sequence 20, Appli
42	321	2.2	2294	4	US-09-854-856-50	Sequence 50, Appli
43	321	2.2	2354	4	US-09-854-856-18	Sequence 18, Appli
44	315.5	2.2	1596	4	US-08-978-277A-4	Sequence 4, Appli
45	313.5	2.2	1863	1	US-08-425-061-16	Sequence 16, Appli

ALIGNMENTS

RESULT 1  
US-07-741-940-2  
; Sequence 2, Application US/07741940  
; Patent No. S352775  
; GENERAL INFORMATION:  
; APPLICANT: ALBERTSEN, HANS  
; APPLICANT: ANAND, RAKESH  
; APPLICANT: CARLSON, MARY  
; APPLICANT: GORDEN, JOANNA  
; APPLICANT: HEDGE, PHILIP J.  
; APPLICANT: JOSLYN, GEOFF  
; APPLICANT: KINZLER, KENNETH  
; APPLICANT: MARKHAM, ALEXANDER F.  
; APPLICANT: NAKAMURA, YUSUKE  
; APPLICANT: THLIVERIS, ANDREW  
; TITLE OF INVENTION: INHERITED AND SOMATIC MUTATIONS OF APC  
; TITLE OF INVENTION: GENE IN COLORECTAL CANCER IN HUMANS  
; NUMBER OF SEQUENCES: 94  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Banner, Birch, McKie & Beckett  
; STREET: 1001 G Street, NW  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20001-4598  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07741,940  
; FILING DATE: 19920109  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Kagan, Sarah A.  
; REGISTRATION NUMBER: 32,141  
; REFERENCE/DOCKET NUMBER: 1107.035574  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-508-9100  
; TELEFAX: 202-508-9299  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2843 amino acids  
; TYPE: AMINO ACID  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-07-741-940-2  
Query Match 100.0%; Score 14575; DB 1; Length 2843;

Best Local Similarity 100.0%; Pred. No. 0;			
Matches 2843; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
Qy	1	MAAASYDQLLKQVAKVENSINLQBELDQNSNHLTKLETASNNKVEVLKOLQSSIDEAM	60
Db	1	MAAASYDQLLKQVAKVENSINLQBELDQNSNHLTKLETASNNKVEVLKOLQSSIDEAM	60
Qy	61	ASSGQIDLLERLKEINLDDSNFPGVKLRSMKLSYSGRSVSSRSRSGECSVPVMSGFPR	120
Db	61	ASSGQIDLLERLKEINLDDSNFPGVKLRSMKLSYSGRSVSSRSRSGECSVPVMSGFPR	120
Qy	121	RGFVNGSRSTGYLEELKERSLLADLXKEEKKWYQAQLQNLTKRIDSLPLTENFSL	180
Db	121	RGFVNGSRSTGYLEELKERSLLADLXKEEKKWYQAQLQNLTKRIDSLPLTENFSL	180
Qy	181	QTDLTRRQLEYEARQIRVAMEEQIGTQDMKRAQRRIARIQIQIEKDILAIRQLQSOAT	240
Db	181	QTDLTRRQLEYEARQIRVAMEEQIGTQDMKRAQRRIARIQIQIEKDILAIRQLQSOAT	240
Qy	241	EAERSQNKHETGSHDAERONEGGVGEINMATSNGCGQSTTRMDHETASVLSSTHSA	300
Db	241	EAERSQNKHETGSHDAERONEGGVGEINMATSNGCGQSTTRMDHETASVLSSTHSA	300
Qy	301	PRRLTSHLGTVEVYSLMLGTHDXDDMSRTLLAMSSQDSCIMRQSGCLPLLIQLL	360
Db	301	PRRLTSHLGTVEVYSLMLGTHDXDDMSRTLLAMSSQDSCIMRQSGCLPLLIQLL	360
Qy	361	HCNDKDSVLLNGSRGSEARASAAALHNIHSQPDQKRGRRIRVLHLLLEQIRAYCETC	420
Db	361	HCNDKDSVLLNGSRGSEARASAAALHNIHSQPDQKRGRRIRVLHLLLEQIRAYCETC	420
Qy	421	WEOEAEHPGMDQKNPMPAPVEHQICPAVCVLMKLSFDEHRHAMNELGGLQAIARELQ	480
Db	421	WEOEAEHPGMDQKNPMPAPVEHQICPAVCVLMKLSFDEHRHAMNELGGLQAIARELQ	480
Qy	481	VDCMYGLTNDHYITLRRVAGMALTNLTGCDVANKATLCSMKGCMALVAQLKSEEDL	540
Db	481	VDCMYGLTNDHYITLRRVAGMALTNLTGCDVANKATLCSMKGCMALVAQLKSEEDL	540
Qy	541	QOVIASVLRNLSWRADVNSKTLREVGSVKALMECALEVKESTLSVLSALNLSAHT	600
Db	541	QOVIASVLRNLSWRADVNSKTLREVGSVKALMECALEVKESTLSVLSALNLSAHT	600
Qy	601	ENKADI CAVDGALAFVGLTYRSQTNLAIIESGGGILRNVSLLIATNEDHRCQILRENN	660
Db	601	ENKADI CAVDGALAFVGLTYRSQTNLAIIESGGGILRNVSLLIATNEDHRCQILRENN	660
Qy	661	CLQTLQLKSHSLTIVSNACGLTNLNSARNPKDQEALWDMGAVSMLKNIHSHKHMIAM	720
Db	661	CLQTLQLKSHSLTIVSNACGLTNLNSARNPKDQEALWDMGAVSMLKNIHSHKHMIAM	720
Qy	721	GSAALRNLANRPAKYKDANIKSPGSLPSLHVKKOKALEAELDAQHLSSETFDNDLS	780
Db	721	GSAALRNLANRPAKYKDANIKSPGSLPSLHVKKOKALEAELDAQHLSSETFDNDLS	780
Qy	781	PKASHRSKQRHQSGLYGDYVFTNRHDDNRSDNFNTGNTVLPYLNTTVLPSSSSRGS	840
Db	781	PKASHRSKQRHQSGLYGDYVFTNRHDDNRSDNFNTGNTVLPYLNTTVLPSSSSRGS	840
Qy	841	LDSSRSKQSLREERGIGLGNVHPATENGTSSKGLQISTTAAQIAKMEVSIAIHTS	900
Db	841	LDSSRSKQSLREERGIGLGNVHPATENGTSSKGLQISTTAAQIAKMEVSIAIHTS	900
Qy	901	QEDRSSGSTTELHCVTDERNALRSSAAHSTHNTYNTFKSENENRITCSMPYAKLEYKRS	960
Db	901	QEDRSSGSTTELHCVTDERNALRSSAAHSTHNTYNTFKSENENRITCSMPYAKLEYKRS	960
Qy	961	NDSLNSVSSNDGYKRGQMKPSESIESYEDDESFKCSYGQYPADLAHKHSANHMDNDGE	1020
Db	961	NDSLNSVSSNDGYKRGQMKPSESIESYEDDESFKCSYGQYPADLAHKHSANHMDNDGE	1020
Qy	1021	LDTPIYNSLKYSDQLNSGRQSPQNERWAPKHIIIEDEIKQSEQRQSRNQSTTVPVYTE	1080

QY 2161 SNKGPRIKPKGKSTLETKKIESEKIGKGGKVKVKSILITGVRNSSEISGOMKQLOAN 2220  
Db 2161 SNKGPRIKPKGKSTLETKKIESEKIGKGGKVKVKSILITGVRNSSEISGOMKQLOAN 2220  
QY 2221 MPSISBGRMTIHPGVNRSSSTSPVSKGPKLPTPKSPSGQATTPSPGAKPSVKS 2280  
Db 2221 MPSISBGRMTIHPGVNRSSSTSPVSKGPKLPTPKSPSGQATTPSPGAKPSVKS 2280  
QY 2281 ELSPVARQTSQIGGSKAPSRGSRDSTPSRPAQPLSRPIQSPGRNISTPGRNGISPPN 2340  
Db 2281 ELSPVARQTSQIGGSKAPSRGSRDSTPSRPAQPLSRPIQSPGRNISTPGRNGISPPN 2340  
QY 2341 KLSQLPRTSPSTASTKSSGSKMSYTSPPGRQMSQONLTQKGLSKNASSIPRSESASKG 2400  
Db 2341 KLSQLPRTSPSTASTKSSGSKMSYTSPPGRQMSQONLTQKGLSKNASSIPRSESASKG 2400  
QY 2401 LQNMNNGANKKVELSRMSSYTSKSGESDRERPVLRQSTFIKEAPSPILRRKLEBSA 2460  
Db 2401 LQNMNNGANKKVELSRMSSYTSKSGESDRERPVLRQSTFIKEAPSPILRRKLEBSA 2460  
QY 2461 SPESISPSRRPASPTRSQAQTPVLSPLPDMSLSTHSSVQAGGWRKLPNLSPTTIEYNDG 2520  
Db 2461 SPESISPSRRPASPTRSQAQTPVLSPLPDMSLSTHSSVQAGGWRKLPNLSPTTIEYNDG 2520  
QY 2521 RPAKHDIARSHSESPSRIPINRSGTWKREHSHKSSILPRVSTWRTGSSSILSASSES 2580  
Db 2521 RPAKHDIARSHSESPSRIPINRSGTWKREHSHKSSILPRVSTWRTGSSSILSASSES 2580  
QY 2581 SEKAKSEDEKHVNSISGTKQSKENOVSAKGTWRKIKENEFSTNISTQTSVSGATNGAES 2640  
Db 2581 SEKAKSEDEKHVNSISGTKQSKENOVSAKGTWRKIKENEFSTNISTQTSVSGATNGAES 2640  
QY 2641 KTLIQMAPAVAKTEDVWRIEDCPINPRSGRSTGNTPTPVVIDSVSEKAMPNIKSDKN 2700  
Db 2641 KTLIQMAPAVAKTEDVWRIEDCPINPRSGRSTGNTPTPVVIDSVSEKAMPNIKSDKN 2700  
QY 2701 QAKQNVGSGVPMRTWGLENRILTSTIQVDAPQKTEIKPGONNPVPVSETNESPIVERT 2760  
Db 2701 QAKQNVGSGVPMRTWGLENRILTSTIQVDAPQKTEIKPGONNPVPVSETNESPIVERT 2760  
QY 2761 PFSSSSSKHSPSGTGAARVTPFNPNPSRKSSADTSARPSQIPTPVNNTKKRDSKT 2820  
Db 2761 PFSSSSSKHSPSGTGAARVTPFNPNPSRKSSADTSARPSQIPTPVNNTKKRDSKT 2820  
QY 2821 DSTESSGTQSPKRGHSGSVLTVSV 2843  
Db 2821 DSTESSGTQSPKRGHSGSVLTVSV 2843

RESULT 2  
US-08-289-548A-2  
Sequence 2, Application US/08289548A  
Patent No. 5648212  
GENERAL INFORMATION:  
APPLICANT: ALBERTSEN, HANS  
APPLICANT: ANAND, RAKESH  
APPLICANT: CARLSON, MARY  
APPLICANT: GRODEN, JOANNA  
APPLICANT: HEDGE, PHILIP J.  
APPLICANT: JOSLYN, GEOFF  
APPLICANT: KINZLER, KENNETH  
APPLICANT: MARKHAM, ALEXANDER F.  
APPLICANT: NAKAMURA, YUSUKE  
APPLICANT: TILIVERIS, ANDREW  
TITLE OF INVENTION: INHERITED AND SOMATIC MUTATIONS OF APC  
NUMBER OF SEQUENCES: 102  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Banner & Allegretti, LTD  
STREET: 1001 G Street, NW  
CITY: Washington  
STATE: D.C.

COUNTRY: USA  
ZIP: 20001-4598  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/289,548A  
FILING DATE: 12-AUG-1994  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Kagan, Sarah A.  
REGISTRATION NUMBER: 32,141  
REFERENCE/DOCKET NUMBER: 1107.46943  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-508-9100  
TELEFAX: 202-508-9239  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2843 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-289-548A-2

Query Match 100.0%; Score 14575; DB 1; Length 2843;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2843; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MAASVDQLLKQVEALKWENSLRQELNEDNSNHLTKLETSANMKVYLKQCSIEDEAM 60  
Db 1 MAASVDQLLKQVEALKWENSLRQELNEDNSNHLTKLETSANMKVYLKQCSIEDEAM 60  
QY 61 ASSQIDLLERLKEINLDSSNFPQVKLRKMSLRYSRGSSVSSRGSGECPVPMGSPFR 120  
Db 61 ASSQIDLLERLKEINLDSSNFPQVKLRKMSLRYSRGSSVSSRGSGECPVPMGSPFR 120  
QY 121 RGFVNGSRESTGYLEELKERSLLADLDKEEKEKQWYLAQLONLTKRISLPLTENFSL 180  
Db 121 RGFVNGSRESTGYLEELKERSLLADLDKEEKEKQWYLAQLONLTKRISLPLTENFSL 180  
QY 181 QDLTRQLRYEARQIRVAMEEQLTGCDMEKQAQRRIARIQOIEKDILRIQLLOSOAT 240  
Db 181 QDLTRQLRYEARQIRVAMEEQLTGCDMEKQAQRRIARIQOIEKDILRIQLLOSOAT 240  
QY 241 EAERSQNHETGSHDAERQNEGGVGEINMATSGNGQSGTTRMDHETASVLSSTHSA 300  
Db 241 EAERSQNHETGSHDAERQNEGGVGEINMATSGNGQSGTTRMDHETASVLSSTHSA 300  
QY 301 PRRLTSHLGTKEVMYSLLSMLGTHDKDDMSRTLLAMSSODSCI SMRQSGCLPLLIQLL 360  
Db 301 PRRLTSHLGTKEVMYSLLSMLGTHDKDDMSRTLLAMSSODSCI SMRQSGCLPLLIQLL 360  
QY 361 HGNDKDSVLLGNRSGSKEARASAAALHNIHSGQDDPKRGRREIRVHLLEQIRAYCETC 420  
Db 361 HGNDKDSVLLGNRSGSKEARASAAALHNIHSGQDDPKRGRREIRVHLLEQIRAYCETC 420  
QY 421 WEWQEAHEPGMDQDKMPMPAPVEHQICPAVCVLMKLSFDEEHRHANNELGGLQAIALLQ 480  
Db 421 WEWQEAHEPGMDQDKMPMPAPVEHQICPAVCVLMKLSFDEEHRHANNELGGLQAIALLQ 480  
QY 481 VDCEMYGLTNDHYSITLRRYAGVALINLTFDGVANKATILCSMGKCMRALVAQKSESEDL 540  
Db 481 VDCEMYGLTNDHYSITLRRYAGVALINLTFDGVANKATILCSMGKCMRALVAQKSESEDL 540  
QY 541 QQVIASVLRNLSWRADVNSKKTIREVGSVKALMECALEVKKESTLKSLSALNLSAHT 600  
Db 541 QQVIASVLRNLSWRADVNSKKTIREVGSVKALMECALEVKKESTLKSLSALNLSAHT 600  
QY 601 ENKADICAVDGAFLVGLTITYSQNTNLAIESGGILNVSLSIATNEDHRCQILRENN 660  
Db 601 ENKADICAVDGAFLVGLTITYSQNTNLAIESGGILNVSLSIATNEDHRCQILRENN 660





```

Db      2821  DSTSSGQTSPKRHSGSYLVTSV  2843
|||||
RESULT 3
US-08-452-654-2
; Sequence 2, Application US/08452654
; Patent No. 5691454
; GENERAL INFORMATION:
; APPLICANT: ALBERTSEN, HANS
; APPLICANT: ANAND, RAKESH
; APPLICANT: CARLSON, MARY
; APPLICANT: GRODEN, JOANNA
; APPLICANT: HEDGE, PHILIP J.
; APPLICANT: JOSLYN, GEOFF
; APPLICANT: KINZLER, KENNETH
; APPLICANT: MARKHAM, ALEXANDER F.
; APPLICANT: NAKAMURA, YUSUKE
; APPLICANT: THLIVERIS, ANDREW
; TITLE OF INVENTION: INHERITED AND SOMATIC MUTATIONS OF APC
; IN THE GENOME OF THE COLON AND RECTAL CANCER IN HUMANS
; NUMBER OF SEQUENCES: 94
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Banner, Birch, McKie & Beckett
; STREET: 1001 G Street, NW
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20001-4598
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/452,654
; FILING DATE: 25-MAY-1995
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/741,940
; FILING DATE: 08-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Kagan, Sarah A.
; REGISTRATION NUMBER: 32,141
; REFERENCE/DOCKET NUMBER: 1107.035574
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-508-9100
; TELEFAX: 202-508-9299
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2843 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: Protein
US-08-452-654-2

Query Match      100.0%; Score 14575; DB 1; Length 2843;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2843; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1  MAASVDQLLKQVEALKVNSLNQLQELSDNSHLTKLETPASNMKVKLQLOGSTEDAM 60
Db      1  MAASVDQLLKQVEALKVNSLNQLQELSDNSHLTKLETPASNMKVKLQLOGSTEDAM 60
QY      61  ASSGQIDLLERLKEALNDSSNFPFGVKLRKSVLSRSGREGSVSSRSGCSVPVNGSPFR 120
Db      61  ASSGQIDLLERLKEALNDSSNFPFGVKLRKSVLSRSGREGSVSSRSGCSVPVNGSPFR 120
QY      121  RGFVNGSRESTGYLEELKESILLADLDKEEKDWMYQAQLNLPKTRIDSLPTENPSL 180
Db      121  RGFVNGSRESTGYLEELKESILLADLDKEEKDWMYQAQLNLPKTRIDSLPTENPSL 180
QY      181  QTDLTROLEYEAQIRVAMEEOLGTCODMKEKAORRIARIOOIEKDILIRIRIOLLOSOAT 240

```

Db 1361 TYCEDVTPICFSCRCSSLSLSSAEDEICGNCQTTQEQADSANTLOIAIEIKGKIGTRSAEDPV 1320  
Qy 1321 SEVPAVCHPRTKSSRIQSSLSSESARHKAVERFPSCAKSPSKGAGTQPKSPPEHYVQET 1380  
Db 1321 SEVPAVCHPRTKSSRIQSSLSSESARHKAVERFPSCAKSPSKGAGTQPKSPPEHYVQET 1380  
Qy 1381 PLMESRCTSVSSLDSESRSTASSVQSEPCSGVUSGIIISPSDLDPSPGQTMPPSRSKTPP 1440  
Db 1381 PLMESRCTSVSSLDSESRSTASSVQSEPCSGVUSGIIISPSDLDPSPGQTMPPSRSKTPP 1440  
Qy 1441 PPPQTAQTKREVKNKAPTAKEKRESGPKQAANAAVQVQLPDADTLLHFATESPDGF 1500  
Db 1441 PPPQTAQTKREVKNKAPTAKEKRESGPKQAANAAVQVQLPDADTLLHFATESPDGF 1500  
Qy 1501 SCSSLSALSILDEPFIOKVELRIMPVQENDNGNETSEQPKESNENQKEAEKIDSE 1560  
Db 1501 SCSSLSALSILDEPFIOKVELRIMPVQENDNGNETSEQPKESNENQKEAEKIDSE 1560  
Qy 1561 KOLLDDDDDDIIELEECIIISAMPTKSSRGKKPAQATASKLPPPVARKPSQLPVYKLLPS 1620  
Db 1561 KOLLDDDDDDIIELEECIIISAMPTKSSRGKKPAQATASKLPPPVARKPSQLPVYKLLPS 1620  
Qy 1621 QNRLOPKHVSFTPGDDMPVYCVGEGTPIINFSTATSLSDLTIESPPNELAAGSVRGGAQ 1680  
Db 1621 QNRLOPKHVSFTPGDDMPVYCVGEGTPIINFSTATSLSDLTIESPPNELAAGSVRGGAQ 1680  
Qy 1681 SGEFEKDDTIPTEGRSDEAGGKTSSVTIPELDDNKAEGDILAEICINAMPKGSHKP 1740  
Db 1681 SGEFEKDDTIPTEGRSDEAGGKTSSVTIPELDDNKAEGDILAEICINAMPKGSHKP 1740  
Qy 1741 PRVKKIMDVOQASASSAPKNOLDGKKKKTPSPVKPIPQNTYRTRVRKNADSKNNLN 1800  
Db 1741 PRVKKIMDVOQASASSAPKNOLDGKKKKTPSPVKPIPQNTYRTRVRKNADSKNNLN 1800  
Qy 1801 AERFVSNDKSKONLKNNSKDNKLPNNEDRVGSGFADSPHHYTPIGTYPVCFSRND 1860  
Db 1801 AERFVSNDKSKONLKNNSKDNKLPNNEDRVGSGFADSPHHYTPIGTYPVCFSRND 1860  
Qy 1861 SLSDLEDDDDVLSREKAEILKAKENKESAEKVTSHTELTSNQQSANKTQAIKQFINR 1920  
Db 1861 SLSDLEDDDDVLSREKAEILKAKENKESAEKVTSHTELTSNQQSANKTQAIKQFINR 1920  
Qy 1921 QOPPILOKQSTFQSSKDIIDRGAATDEKLOFAIENTVCFSHNSLSLSDIDQENN 1980  
Db 1921 QOPPILOKQSTFQSSKDIIDRGAATDEKLOFAIENTVCFSHNSLSLSDIDQENN 1980  
Qy 1981 NKNEPIKETEPDPSQGEPSKQASGAPKSFHVEDTPVCFSRNSLSLSDIDSEDDLLQ 2040  
Db 1981 NKNEPIKETEPDPSQGEPSKQASGAPKSFHVEDTPVCFSRNSLSLSDIDSEDDLLQ 2040  
Qy 2041 ECISAMPKPKKRLKDNKESKPRNMGILGEBDLTLDKIQRPDSEHGLSPDSENF 2100  
Db 2041 ECISAMPKPKKRLKDNKESKPRNMGILGEBDLTLDKIQRPDSEHGLSPDSENF 2100  
Qy 2101 WKALQEGANSIVSLHQAACLSRQASDSDSILSKSGISGLGPPFHLTPQEEKPT 2160  
Db 2101 WKALQEGANSIVSLHQAACLSRQASDSDSILSKSGISGLGPPFHLTPQEEKPT 2160  
Qy 2161 SNKGPRILKPEKSTLTETKIESKGIKGGKVKYKSLITGKVRNSNIEISGQWKQPLQAN 2220  
Db 2161 SNKGPRILKPEKSTLTETKIESKGIKGGKVKYKSLITGKVRNSNIEISGQWKQPLQAN 2220  
Qy 2221 MPSTSRGRTMIHICVNRNSSSSTSPVSKGPPKLTTPASKPSSEGQATATTPRGAKEPVSKS 2280  
Db 2221 MPSTSRGRTMIHICVNRNSSSSTSPVSKGPPKLTTPASKPSSEGQATATTPRGAKEPVSKS 2280  
Qy 2281 ELSFVARQTSIQGSSKAPSGSRDSTPSRPAQOPLSRPIQSPGRNSISPGRNGISPPN 2340  
Db 2281 ELSFVARQTSIQGSSKAPSGSRDSTPSRPAQOPLSRPIQSPGRNSISPGRNGISPPN 2340  
Qy 2341 KLSQLPRTSSPSTASTSSGSKNSYTSRGRNSQOQLTKOTGLSKNASSIPRSESASKG 2400  
Db 2341 KLSQLPRTSSPSTASTSSGSKNSYTSRGRNSQOQLTKOTGLSKNASSIPRSESASKG 2400

Qy 2401 LNMNMNGANKKVELSRMSTKSSGSEDRSERPVLVRQSTFIKEAPSPTLRKLBSA 2460  
Db 2401 LNMNMNGANKKVELSRMSTKSSGSEDRSERPVLVRQSTFIKEAPSPTLRKLBSA 2460  
Qy 2461 SPESLSPSSRPASPTRSQAQTFVLSPSLPDMSLSTHSSVQAGGWRKLPNLSPTIEYNDG 2520  
Db 2461 SPESLSPSSRPASPTRSQAQTFVLSPSLPDMSLSTHSSVQAGGWRKLPNLSPTIEYNDG 2520  
Qy 2521 RPAKHDIARSHSPSPSLPNNRSGTWKREHSHSSSLPRVSTWRTGSSSSILSASES 2580  
Db 2521 RPAKHDIARSHSPSPSLPNNRSGTWKREHSHSSSLPRVSTWRTGSSSSILSASES 2580  
Qy 2581 SEKAKSEDEKHVNSISGTKQSKENQVSAKGTWRKIKENEPSPTNSTQTVSSSGATNGAES 2640  
Db 2581 SEKAKSEDEKHVNSISGTKQSKENQVSAKGTWRKIKENEPSPTNSTQTVSSSGATNGAES 2640  
Qy 2641 KTLIYQMAPAVSKTEDVWVRIEDCPINNPRSGRPTGNTPPVIDSVSEKANPNIKDSKN 2700  
Db 2641 KTLIYQMAPAVSKTEDVWVRIEDCPINNPRSGRPTGNTPPVIDSVSEKANPNIKDSKN 2700  
Qy 2701 QAKQNVGNGSVPMRTVGLNRLTSFIQVADPQKGTIKQGNPNPVPVSETNESPIVERT 2760  
Db 2701 QAKQNVGNGSVPMRTVGLNRLTSFIQVADPQKGTIKQGNPNPVPVSETNESPIVERT 2760  
Qy 2761 PFSSSSSKSHSPSGTVAARVTPFNYPNPSRKSSADSTARSPOIPTPVNNNTKKRDSKT 2820  
Db 2761 PFSSSSSKSHSPSGTVAARVTPFNYPNPSRKSSADSTARSPOIPTPVNNNTKKRDSKT 2820  
Qy 2821 DSTESSGTSQPKRHSGSYLVTSV 2843  
Db 2821 DSTESSGTSQPKRHSGSYLVTSV 2843

RESULT 4  
US-08-370-235A-2  
; Sequence 2, Application US/08370235A  
; Patent No. 5910418  
; GENERAL INFORMATION:  
; APPLICANT: VOGELSTEIN, BERT  
; APPLICANT: KINZLER, KENNETH W.  
; APPLICANT: HILL, DAVID E.  
; APPLICANT: JOHNSON, KAREN A.  
; TITLE OF INVENTION: ANTIBODIES AND ASSAYS FOR DETERMINING  
; TITLE OF INVENTION: MUTATIONS IN THE APC GENE  
; NUMBER OF SEQUENCES: 4  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: BANNER & WITCOFF, LTD.  
; STREET: 1001 G STREET, N.W.  
; CITY: WASHINGTON  
; STATE: DC  
; COUNTRY: US  
; ZIP: 20001  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/370,235A  
; FILING DATE: 01-JAN-1995  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: KAGAN, SARAH A.  
; REGISTRATION NUMBER: 32,141  
; REFERENCE/DOCKET NUMBER: 01107.48688  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202 508 9100  
; TELEFAX: 202 508 9299  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2843 amino acids  
; TYPE: amino acid

; TOPOLOGY: linear									
; MOLECULE TYPE: protein									
US-08-370-235A-2									
Query Match 100.0%; Score 14575; DB 2; Length 2843;									
Best Local Similarity 100.0%; Pred. No. 0;									
Matches 2843; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
QY	1	MAAASYDQLLKQVEALKWENLNRLQELNEDNHNLTQLETSANMKEVLLKQLQGSIEDEAM	60						
DB	1	MAAASYDQLLKQVEALKWENLNRLQELNEDNHNLTQLETSANMKEVLLKQLQGSIEDEAM	60						
QY	61	ASSGOIDILRLKEILNLDNFPNPGVKLSKYSLSYSGREGSVSRSGECSVPWMSGFPR	120						
DB	61	ASSGOIDILRLKEILNLDNFPNPGVKLSKYSLSYSGREGSVSRSGECSVPWMSGFPR	120						
QY	121	RGFVNGSRESTGYLLEEKERSLLIADLDKEEKDWYIAQNLTKRIDSPLPTENPFL	180						
DB	121	RGFVNGSRESTGYLLEEKERSLLIADLDKEEKDWYIAQNLTKRIDSPLPTENPFL	180						
QY	181	QDILTRRQLEYEARQIRVAMEEQLGTQDMKRAQRRIARIQOIEKDLIRIQLQSQAT	240						
DB	181	QDILTRRQLEYEARQIRVAMEEQLGTQDMKRAQRRIARIQOIEKDLIRIQLQSQAT	240						
QY	241	EAERSQNKHTGSHDARQNEGQVGEINMATSGNGOGCTTRMDHETASVLSSSSTHSA	300						
DB	241	EAERSQNKHTGSHDARQNEGQVGEINMATSGNGOGCTTRMDHETASVLSSSSTHSA	300						
QY	301	PRRLTSHLGTKEVWYLSLLMLGTHDXDDMGRTLLAMSSQDSCISMRQSCCLPLLIQLL	360						
DB	301	PRRLTSHLGTKEVWYLSLLMLGTHDXDDMGRTLLAMSSQDSCISMRQSCCLPLLIQLL	360						
QY	361	HGNDRKSVLLNGSRGSKARAPASAAAHNIITHSQPDDKRGREIRVLHLLEIQIRAYCETC	420						
DB	361	HGNDRKSVLLNGSRGSKARAPASAAAHNIITHSQPDDKRGREIRVLHLLEIQIRAYCETC	420						
QY	421	NEWQEAHEPGMDQDNMPAPVEHQICPAVCVLMKLSFDEEHRHAMNELGLQIAELQL	480						
DB	421	NEWQEAHEPGMDQDNMPAPVEHQICPAVCVLMKLSFDEEHRHAMNELGLQIAELQL	480						
QY	481	VDCEMYGLTNDHYSITLRYAGMALTNITFGDVANKATILCSMKGCMRALVAQLKSEDL	540						
DB	481	VDCEMYGLTNDHYSITLRYAGMALTNITFGDVANKATILCSMKGCMRALVAQLKSEDL	540						
QY	541	QOVIASVLNLSRADVNSKTLREVGSVKALMECALEVKKESTLKSVLNALWLSAHC	600						
DB	541	QOVIASVLNLSRADVNSKTLREVGSVKALMECALEVKKESTLKSVLNALWLSAHC	600						
QY	601	ENKADICAVDGAFLVGLTYRSQNTLAIIESGGILNRVSSLIATNEDHRQILRENN	660						
DB	601	ENKADICAVDGAFLVGLTYRSQNTLAIIESGGILNRVSSLIATNEDHRQILRENN	660						
QY	661	CLQTLQHLKSHSLTIVSNACGTLNLSARNPKDQDALWDMGAVSMKLNLIHSHKMIAM	720						
DB	661	CLQTLQHLKSHSLTIVSNACGTLNLSARNPKDQDALWDMGAVSMKLNLIHSHKMIAM	720						
QY	721	GSAALRNLMANRPARYKDANIMSPGSSILPSLHVVRKQKALEDAQLHSETFDNIDLS	780						
DB	721	GSAALRNLMANRPARYKDANIMSPGSSILPSLHVVRKQKALEDAQLHSETFDNIDLS	780						
QY	781	PKASHRSKORHKQSLYGDVFDNTRHDNRSDNFTGNMTVLSPLYNTLTVLPSSSSSRGS	840						
DB	781	PKASHRSKORHKQSLYGDVFDNTRHDNRSDNFTGNMTVLSPLYNTLTVLPSSSSSRGS	840						
QY	841	LDSSRSKDRSLERERIGLGNVHPATENPGTSSKRGQLISTTAAQIAKWEEVSAIHTS	900						
DB	841	LDSSRSKDRSLERERIGLGNVHPATENPGTSSKRGQLISTTAAQIAKWEEVSAIHTS	900						
QY	901	QEDRSGSTTELHCVTDERNALRRSAAHTNTYNTFKSENSNRTCSMPYAKLEYKRS	960						
DB	901	QEDRSGSTTELHCVTDERNALRRSAAHTNTYNTFKSENSNRTCSMPYAKLEYKRS	960						
QY	961	NDLSNSVSSNDGYRGKQMKPSIESYSEDDSKFCSYGQYPADLAHKIHSANHMDNDGE	1020						

DB	961	NDLSNSVSSNDGYRGKQMKPSIESYSEDDSKFCSYGQYPADLAHKIHSANHMDNDGE	1020						
QY	1021	LDPINYSKYDEQLNSGRQSPQSNRWARPKHIEDEIKQSEQRQSNQSTTYPVYTE	1080						
DB	1021	LDPINYSKYDEQLNSGRQSPQSNRWARPKHIEDEIKQSEQRQSNQSTTYPVYTE	1080						
QY	1081	STDDKHLKQPHFGQCEVSPYRSRGANGSETNRVSGNHNQINQVQSILCOEDDIEDDKP	1140						
DB	1081	STDDKHLKQPHFGQCEVSPYRSRGANGSETNRVSGNHNQINQVQSILCOEDDIEDDKP	1140						
QY	1141	TNYSERYSEEOHEBERPTNYSIKYNEEKRVHDQPIDYSLKYATDIPSSQKQSFSSKS	1200						
DB	1141	TNYSERYSEEOHEBERPTNYSIKYNEEKRVHDQPIDYSLKYATDIPSSQKQSFSSKS	1200						
QY	1201	SSQSKTEHMGSSSENTSTPSNAKRONQLHPSSQAQSRSGQPOKAATCKVSSINOETIQ	1260						
DB	1201	SSQSKTEHMGSSSENTSTPSNAKRONQLHPSSQAQSRSGQPOKAATCKVSSINOETIQ	1260						
QY	1261	TYCVDTPICFRCSSLSLSAEDSIGCQTQEQADSANTLIQIAIKGIGTRSAEDPV	1320						
DB	1261	TYCVDTPICFRCSSLSLSAEDSIGCQTQEQADSANTLIQIAIKGIGTRSAEDPV	1320						
QY	1321	SEVPAYSQHPRTKSSRLQSSLSSESARHKAVEFPSCAKSPKSGAQTPKSPEHYVQET	1380						
DB	1321	SEVPAYSQHPRTKSSRLQSSLSSESARHKAVEFPSCAKSPKSGAQTPKSPEHYVQET	1380						
QY	1381	PLMFSCTSVSSLDSPESRSIASSVQSEPCSGMVSGIISPSDLDPSPGOTMPPSRSKTTP	1440						
DB	1381	PLMFSCTSVSSLDSPESRSIASSVQSEPCSGMVSGIISPSDLDPSPGOTMPPSRSKTTP	1440						
QY	1441	PPQTAQTKREVFNKAPTAETKESGPKQAANAAVQVRQVLPDADTLHLHFATESPDGF	1500						
DB	1441	PPQTAQTKREVFNKAPTAETKESGPKQAANAAVQVRQVLPDADTLHLHFATESPDGF	1500						
QY	1501	SCSSSLSALSLDPPPTOKVLEIRIMPVOENDNGNETESEQPKESNEKEAEKIIDSE	1560						
DB	1501	SCSSSLSALSLDPPPTOKVLEIRIMPVOENDNGNETESEQPKESNEKEAEKIIDSE	1560						
QY	1561	KDILLDDSDDDIIEBECIISAMPTKSSRKGKPAQATASCLPPPVARKPSQSPVYKLLPS	1620						
DB	1561	KDILLDDSDDDIIEBECIISAMPTKSSRKGKPAQATASCLPPPVARKPSQSPVYKLLPS	1620						
QY	1621	QNRLOPKQVHSFTPGDDMPRVYCVETGPIINFSTATSLSDLTIESPPNELAAGVGVRGAQ	1680						
DB	1621	QNRLOPKQVHSFTPGDDMPRVYCVETGPIINFSTATSLSDLTIESPPNELAAGVGVRGAQ	1680						
QY	1681	SGFEKEDTITPTEGRSTDEAQQGKTSVTIPELDDNKAEBGDILACINSAMPKGSHP	1740						
DB	1681	SGFEKEDTITPTEGRSTDEAQQGKTSVTIPELDDNKAEBGDILACINSAMPKGSHP	1740						
QY	1741	FRVKIMDQVQOASASSAPNKNQLDGKKKPTSPVKPIPQNTYTRVRKNADSKNNLN	1800						
DB	1741	FRVKIMDQVQOASASSAPNKNQLDGKKKPTSPVKPIPQNTYTRVRKNADSKNNLN	1800						
QY	1801	AERFSDNKSQKQNLKNSKDFNDKLPNNEDVRGSFAPDSPHHTPIEGTFCPSRND	1860						
DB	1801	AERFSDNKSQKQNLKNSKDFNDKLPNNEDVRGSFAPDSPHHTPIEGTFCPSRND	1860						
QY	1861	SLGSLDPPDDDDVLSREKAEIRKAKENKESAEKVTSHTELTSNQQSANKTQAIKQPINR	1920						
DB	1861	SLGSLDPPDDDDVLSREKAEIRKAKENKESAEKVTSHTELTSNQQSANKTQAIKQPINR	1920						
QY	1921	GQPKPILQKQSTFPQSSKDIPTDRGAATDEKLQNFALIENTPVCFSHNSSSLSLSDIOENN	1980						
DB	1921	GQPKPILQKQSTFPQSSKDIPTDRGAATDEKLQNFALIENTPVCFSHNSSSLSLSDIOENN	1980						
QY	1981	NKENPTEKETEPDQSGECPKQASGYAPKSFVVEDTPVCFSSNSSLSLSISEDDLLQ	2040						
DB	1981	NKENPTEKETEPDQSGECPKQASGYAPKSFVVEDTPVCFSSNSSLSLSISEDDLLQ	2040						
QY	2041	ECISSAMPKPKPSRLKGDNEKHSRPNMGILGEDTLDKDIQRPDSRGLSPDSNFD	2100						

Db 2041 ECISAMPKKKPSRLKCDNEKHSRNMGGILGEDLTLDLKDIOQPDSEHGLSPDSENF 2100  
 Qy 2101 WKAIQEGANSIVSSHQAACLSRQASDSILSKGISLGSFPHLTPOQEKPT 2160  
 Db 2101 WKAIQEGANSIVSSHQAACLSRQASDSILSKGISLGSFPHLTPOQEKPT 2160  
 Qy 2161 SNKGPRILKPEKSTLTETKTESKGIKGGKVKYKSLITGKVRNSSEISQWKQPLQAN 2220  
 Db 2161 SNKGPRILKPEKSTLTETKTESKGIKGGKVKYKSLITGKVRNSSEISQWKQPLQAN 2220  
 Qy 2221 MPSISRGRTMIHIFGVNRSSTSPVSKGPPPLKTPASKSPSEGQATTTPRGAKPSVKS 2280  
 Db 2221 MPSISRGRTMIHIFGVNRSSTSPVSKGPPPLKTPASKSPSEGQATTTPRGAKPSVKS 2280  
 Qy 2281 ELSPVARTQISQIGSSKAPRSGRDTPRPAQOPLSRPIQSPGRNISIPGRNGISPPN 2340  
 Db 2281 ELSPVARTQISQIGSSKAPRSGRDTPRPAQOPLSRPIQSPGRNISIPGRNGISPPN 2340  
 Qy 2341 KLSQLPRTSSPTASTKSSGSKMYSYTPGRQMSQQNLTKQTGLSKNASSIPRSESASKG 2400  
 Db 2341 KLSQLPRTSSPTASTKSSGSKMYSYTPGRQMSQQNLTKQTGLSKNASSIPRSESASKG 2400  
 Qy 2401 LQWANGNGANKKVELSRMASTKSSGSDRSRPPVLRQSTFIKEAPSTLRKLEESA 2460  
 Db 2401 LQWANGNGANKKVELSRMASTKSSGSDRSRPPVLRQSTFIKEAPSTLRKLEESA 2460  
 Qy 2461 SFESLSPSSRPASPTRSQOATPVLSPLPDMSLSTHSSVQAGWRKLPNLSPTIEYNDG 2520  
 Db 2461 SFESLSPSSRPASPTRSQOATPVLSPLPDMSLSTHSSVQAGWRKLPNLSPTIEYNDG 2520  
 Qy 2521 RPAKRHDIAHSHSPRLPINRSGTWKREHSHSSSLPRVSTWRTTSGSSSILSASSES 2580  
 Db 2521 RPAKRHDIAHSHSPRLPINRSGTWKREHSHSSSLPRVSTWRTTSGSSSILSASSES 2580  
 Qy 2581 SEKAKSDEKHVNSISGTQKSKENOVSAKGTWRKIKENEFSPNTNSTQTVSSGATNGAES 2640  
 Db 2581 SEKAKSDEKHVNSISGTQKSKENOVSAKGTWRKIKENEFSPNTNSTQTVSSGATNGAES 2640  
 Qy 2641 KTLIYQAPAVSKTEDVWVRIEDCPINNRSGRSPGTNTPPVDSVSEKANPNIKDSKN 2700  
 Db 2641 KTLIYQAPAVSKTEDVWVRIEDCPINNRSGRSPGTNTPPVDSVSEKANPNIKDSKN 2700  
 Qy 2701 QAKQNVGSGVPMRTVGLNRLTSFIOVDAPDQKTEIKPQNNPVPVSETNESPIVERT 2760  
 Db 2701 QAKQNVGSGVPMRTVGLNRLTSFIOVDAPDQKTEIKPQNNPVPVSETNESPIVERT 2760  
 Qy 2761 PFSSSSSSKSSPSGTVAARVTPFNPNPSPRKSSADSTSPRSQIPTPVNNNTKKRDSKT 2820  
 Db 2761 PFSSSSSSKSSPSGTVAARVTPFNPNPSPRKSSADSTSPRSQIPTPVNNNTKKRDSKT 2820  
 Qy 2821 DSTESSGTQSPKRGSGSYLTVS 2843  
 Db 2821 DSTESSGTQSPKRGSGSYLTVS 2843

RESULT 5

US-08-449-731-2

; Sequence 2, Application US/08449731

; Patent No. 6413727

; GENERAL INFORMATION:

; APPLICANT: ALBERTSEN, HANS

; ANAND, RAKESH

; CARLSON, MARY

; GRODEN, JOANNA

; HEDGE, PHILIP J.

; JOSLYN, GEOFF

; KINZLER, KENNETH

; MARKHAM, ALEXANDER F.

; NAKAMURA, YUSUKE

; THLIVERIS, ANDREW

; TITLE OF INVENTION: INHERITED AND SOMATIC MUTATIONS OF APC

; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Banner & Allegretti, LTD  
 ; STREET: 1001 G Street, NW  
 ; CITY: Washington  
 ; STATE: D.C.  
 ; COUNTRY: USA  
 ; ZIP: 20001-4598  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patent In Release #1.0, Version #1.25  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/449,731  
 ; FILING DATE: 25-May-1995  
 ; CLASSIFICATION: <Unknown>  
 ; PRIOR APPLICATION NUMBER: 08/289,548  
 ; FILING DATE: 12-AUG-1994  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Kagan, Sarah A.  
 ; REGISTRATION NUMBER: 32,141  
 ; REFERENCE/DOCKET NUMBER: 1107.46943  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 202-508-9100  
 ; TELEFAX: 202-508-9299  
 ; INFORMATION FOR SEQ ID NO: 2:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 2843 amino acids  
 ; TYPE: amino acid  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: protein  
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 2:  
 ; US-08-449-731-2

Query Match 100.0%; Score 14575; DB 4; Length 2843;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 2843; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAAASYDQLLKQVAKLVKENSINLRQLEDNSNHLTKLETEASNKVELKQLQGSIEDEAM 60  
 Db 1 MAAASYDQLLKQVAKLVKENSINLRQLEDNSNHLTKLETEASNKVELKQLQGSIEDEAM 60  
 Qy 61 ASSQIDLLERKELNLDSSNFPVGLKRSKMSLSYSGRSGSVSSRSGSGSPVPMGSPFR 120  
 Db 61 ASSQIDLLERKELNLDSSNFPVGLKRSKMSLSYSGRSGSVSSRSGSGSPVPMGSPFR 120  
 Qy 121 RGFVNGSRESGYLEELKERSILLADLDKEEKDWYQAQLNLTIRIDSLPTENFSL 180  
 Db 121 RGFVNGSRESGYLEELKERSILLADLDKEEKDWYQAQLNLTIRIDSLPTENFSL 180  
 Qy 181 QTDLTRRQLEYEARQIRVAMEEQLTGTCQDMKRAQRRIARIQIEKILIRIQLLOQAT 240  
 Db 181 QTDLTRRQLEYEARQIRVAMEEQLTGTCQDMKRAQRRIARIQIEKILIRIQLLOQAT 240  
 Qy 241 EAERSSQKHETGSHDAERQNEGQVGEINMATSNGQGGSTTRMDHETASVLSSTHSA 300  
 Db 241 EAERSSQKHETGSHDAERQNEGQVGEINMATSNGQGGSTTRMDHETASVLSSTHSA 300  
 Qy 301 PRRLTSHLGTVMVYSLLSMLGTHDKDMSRTLAMSSQSDSCISMRQSGCPLLIQLL 360  
 Db 301 PRRLTSHLGTVMVYSLLSMLGTHDKDMSRTLAMSSQSDSCISMRQSGCPLLIQLL 360  
 Qy 361 HGNDKDSVLLGNRSGSKARARASAAALHNIHSOPDKRGERIRVHLLEQIYACETC 420  
 Db 361 HGNDKDSVLLGNRSGSKARARASAAALHNIHSOPDKRGERIRVHLLEQIYACETC 420  
 Qy 421 WEWQEAHEPGMDQDKNPNPAPVEHQICPAVCVLMKLSFDEEHRHAMNELGGLQAIALLQ 480  
 Db 421 WEWQEAHEPGMDQDKNPNPAPVEHQICPAVCVLMKLSFDEEHRHAMNELGGLQAIALLQ 480  
 Qy 481 VDCMEYGLTNDHYSITLRRYAGVALTNLTFGDVANKATLCSMKGCWRALVAQLKSEEDL 540

481 VDCMYGUTNDHYSITLRYAGMALTNLTFGVANKATLCSMKOCMALVAQLKSESED 540  
481 QOVIASVLRLNLSWRADVNSKTLREVGSVKALMECALEVKESTLKVSLGALWNLSAHT 600  
481 QOVIASVLRLNLSWRADVNSKTLREVGSVKALMECALEVKESTLKVSLGALWNLSAHT 600  
601 ENKADICAVDGCALAFVGLTYRSGTNTLAIIESGGGILRVNLSLIATNEDHROILRENN 660  
601 ENKADICAVDGCALAFVGLTYRSGTNTLAIIESGGGILRVNLSLIATNEDHROILRENN 660  
661 CLQTLQLHLKSHSLTIVSNACTLWNLARNPKDQEALWDMGVSMKLNLIHSKHKMIAM 720  
661 CLQTLQLHLKSHSLTIVSNACTLWNLARNPKDQEALWDMGVSMKLNLIHSKHKMIAM 720  
721 GSAALRLNLMANRPKAKYKADANIMPSSSLPSLHVYRKQKALEDAOHLSETFNDNLNLS 780  
721 GSAALRLNLMANRPKAKYKADANIMPSSSLPSLHVYRKQKALEDAOHLSETFNDNLNLS 780  
781 PKASHRSKORHKQSLGYDVTNRHDDNRDNFNTGNMTVLSPLYNTTIVLPSSSSSRGS 840  
781 PKASHRSKORHKQSLGYDVTNRHDDNRDNFNTGNMTVLSPLYNTTIVLPSSSSSRGS 840  
841 LDSSRSKDRSLERBRGIGLGNYPHATENPGTSSKRGLOISTTAAQIAKVMEEVSAIHTS 900  
841 LDSSRSKDRSLERBRGIGLGNYPHATENPGTSSKRGLOISTTAAQIAKVMEEVSAIHTS 900  
901 QEDRSSGTTTELHCVTDERNALRRSSAAHTSNNTYNTFKSENSNRTCSMPYAKLEYKRSS 960  
901 QEDRSSGTTTELHCVTDERNALRRSSAAHTSNNTYNTFKSENSNRTCSMPYAKLEYKRSS 960  
961 NDSLNSVSNDDGYKRGQMKPSIESYEDDESKFCYQYQPADLAHKIHSANHMDNDNGE 1020  
961 NDSLNSVSNDDGYKRGQMKPSIESYEDDESKFCYQYQPADLAHKIHSANHMDNDNGE 1020  
1021 LDTPNYSIKYDSEOLNSGRQSPSONERWAPKPHIIEDEIKQSEORQSRNOSTYVPVYTE 1080  
1021 LDTPNYSIKYDSEOLNSGRQSPSONERWAPKPHIIEDEIKQSEORQSRNOSTYVPVYTE 1080  
1081 STDDXHLAFQPHFGQOECVPSYRSGANGSETRVGSNHGINQNVQSLSQCEDDEDYDDKP 1140  
1081 STDDXHLAFQPHFGQOECVPSYRSGANGSETRVGSNHGINQNVQSLSQCEDDEDYDDKP 1140  
1141 TNYSRYSEREEHHEERPTNYSIKYNEEKRHVDOPIDYSLKYATDIPSSOKQSFSEKS 1200  
1141 TNYSRYSEREEHHEERPTNYSIKYNEEKRHVDOPIDYSLKYATDIPSSOKQSFSEKS 1200  
1201 SSGQSKTEHMSSENSTPSSNAKRONQLHPSSAQSRSGOPOKAATCKVSSINQETIQ 1260  
1201 SSGQSKTEHMSSENSTPSSNAKRONQLHPSSAQSRSGOPOKAATCKVSSINQETIQ 1260  
1261 TYCVEDTPICFGRCSLSLSAEDIEGQNTQTQEADSAANTLQIAIEIKGKIGTRSAEDPV 1320  
1261 TYCVEDTPICFGRCSLSLSAEDIEGQNTQTQEADSAANTLQIAIEIKGKIGTRSAEDPV 1320  
1321 SEVPVAVSOHPRTKSRLOGSSLSSESARHKAVERPFGSAKSPSKCAOTPKSPPEHYVOET 1380  
1321 SEVPVAVSOHPRTKSRLOGSSLSSESARHKAVERPFGSAKSPSKCAOTPKSPPEHYVOET 1380  
1381 PLMFRCSTSVSLDPSFESRSIASSVQSEPCSGWVGIIISPSDLPSQGTMPPSRSKTPP 1440  
1381 PLMFRCSTSVSLDPSFESRSIASSVQSEPCSGWVGIIISPSDLPSQGTMPPSRSKTPP 1440  
1441 PPPQTAQTKREVPKNAKPTAEKRESCPOAAVNAVQVLPDADTLHLHATESTPDGF 1500  
1441 PPPQTAQTKREVPKNAKPTAEKRESCPOAAVNAVQVLPDADTLHLHATESTPDGF 1500  
1501 SCSSLSLALSLDBPFIQKDVLRIMPPVQENDNGNETESEQPKSENENQKEAEXTIDSE 1560  
1501 SCSSLSLALSLDBPFIQKDVLRIMPPVQENDNGNETESEQPKSENENQKEAEXTIDSE 1560  
1561 KOLLDDSDDDDIIEILBECIISAMPTKSRKGGKPAQTASKLPPPVARKPSQLPVYKLLPS 1620  
1561 KOLLDDSDDDDIIEILBECIISAMPTKSRKGGKPAQTASKLPPPVARKPSQLPVYKLLPS 1620

1621 QNELQPKHVSTPFGDDMPRVYCVBGTPIINFSTATSLSDLTIESPPNELAAGEVGGQAQ 1680  
1621 QNELQPKHVSTPFGDDMPRVYCVBGTPIINFSTATSLSDLTIESPPNELAAGEVGGQAQ 1680  
1681 SGBFEKRDITPTEGRSTDEAQQGKTSSVTIPELDDNKAEEGDILABECINSAMPKGSHPK 1740  
1681 SGBFEKRDITPTEGRSTDEAQQGKTSSVTIPELDDNKAEEGDILABECINSAMPKGSHPK 1740  
1741 FRVKIMQVQQAASASSAPNKNQOLDGKKKKTSPVKPIPONTETRYTRVRKNADSKNNLN 1800  
1741 FRVKIMQVQQAASASSAPNKNQOLDGKKKKTSPVKPIPONTETRYTRVRKNADSKNNLN 1800  
1801 AERFSDNKKDKONLKNNSKDFNDKLPNNEDRVGSPAFDPGPHHTPIEGTYPYCFERN 1860  
1801 AERFSDNKKDKONLKNNSKDFNDKLPNNEDRVGSPAFDPGPHHTPIEGTYPYCFERN 1860  
1861 SLSLDPDDDDVLSREKAELRAKAKENKSEAKVTSHTELTNQQSANTQAIKQPINR 1920  
1861 SLSLDPDDDDVLSREKAELRAKAKENKSEAKVTSHTELTNQQSANTQAIKQPINR 1920  
1921 GQPKPILQKOSTEPOSQKDIIPDRGAATDEKLPNNEDRVGSPAFDPGPHHTPIEGTYPYCFERN 1980  
1921 GQPKPILQKOSTEPOSQKDIIPDRGAATDEKLPNNEDRVGSPAFDPGPHHTPIEGTYPYCFERN 1980  
1981 NKENEPIKETEPDPSQGFPSKPOASGYAPKSFHVEDTPVCFERNSSLSLSIDSEDDLQ 2040  
1981 NKENEPIKETEPDPSQGFPSKPOASGYAPKSFHVEDTPVCFERNSSLSLSIDSEDDLQ 2040  
2041 ECISAMPKPKSRILKCDNEKHSPRNMGILGEDLTDLKQIQRDPDSEHGLSPDENFD 2100  
2041 ECISAMPKPKSRILKCDNEKHSPRNMGILGEDLTDLKQIQRDPDSEHGLSPDENFD 2100  
2101 WKAIQEGANSIVSSLHQAACCLSRQAASDSDSILSLXSGISLGSPPFHLTPDQEBKPT 2160  
2101 WKAIQEGANSIVSSLHQAACCLSRQAASDSDSILSLXSGISLGSPPFHLTPDQEBKPT 2160  
2161 SNKGPRILKPKGKSTLTETKIESEKIGKGGKVKYKSLITGKVRNSSEISGQWKQLOAN 2220  
2161 SNKGPRILKPKGKSTLTETKIESEKIGKGGKVKYKSLITGKVRNSSEISGQWKQLOAN 2220  
2221 MPSTSRGRTMIHPIGVNNSSTSPVSKGPPPLKTPASKSPSEGOATATSPRGAKEPVKS 2280  
2221 MPSTSRGRTMIHPIGVNNSSTSPVSKGPPPLKTPASKSPSEGOATATSPRGAKEPVKS 2280  
2281 ELSFVARQTSQIGGSKAPSRGSRDSTPSRPAQQLSRPIQSPGRNSISPGRNGISPPN 2340  
2281 ELSFVARQTSQIGGSKAPSRGSRDSTPSRPAQQLSRPIQSPGRNSISPGRNGISPPN 2340  
2341 KLSQLPRTSPSTASTKSSGSGKMSYTSPCRQMSQONLTQOTGLSKXNASSIPSESEASKG 2400  
2341 KLSQLPRTSPSTASTKSSGSGKMSYTSPCRQMSQONLTQOTGLSKXNASSIPSESEASKG 2400  
2401 LNQMNGNGANKKVELSRMSTKSSGSESDRSPVLRQSTFIKEAPSTLRRKLEESA 2460  
2401 LNQMNGNGANKKVELSRMSTKSSGSESDRSPVLRQSTFIKEAPSTLRRKLEESA 2460  
2461 SPSLSPPSPAPSTRSOAQTPVLSPLPMSISTHSSVQAGGWRKLPNLSPTIENDG 2520  
2461 SPSLSPPSPAPSTRSOAQTPVLSPLPMSISTHSSVQAGGWRKLPNLSPTIENDG 2520  
2521 RPAKHDIARSHSPSRPLPNRSKXSHSSSLPRVSTWRTTGSSSSILSASSES 2580  
2521 RPAKHDIARSHSPSRPLPNRSKXSHSSSLPRVSTWRTTGSSSSILSASSES 2580  
2581 SEKAKSEDEKHNVISGOTKQKSNQVSAKWTWKIKENEFSPNTNSTQTVSSGATNGAES 2640  
2581 SEKAKSEDEKHNVISGOTKQKSNQVSAKWTWKIKENEFSPNTNSTQTVSSGATNGAES 2640  
2641 KTLIYQAPAVSKTDEWVRIEDCPINPRSGSPGTGNTPPVIDSVSEKANPNIKDSKDN 2700  
2641 KTLIYQAPAVSKTDEWVRIEDCPINPRSGSPGTGNTPPVIDSVSEKANPNIKDSKDN 2700

QY 2701 QAKQVNGSVPMRTVGLNRLTSFIQVADPDQKGTBIKPGQNNPVPVSVTNSPIVERT 2760  
 Db 2701 QAKQVNGSVPMRTVGLNRLTSFIQVADPDQKGTBIKPGQNNPVPVSVTNSPIVERT 2760  
 QY 2761 PFSSSSSKHSSPSGTVAARVTPNNYNPSPKSSADSTSARPSOIPTPVNNNTKKRDSKT 2820  
 Db 2761 PFSSSSSKHSSPSGTVAARVTPNNYNPSPKSSADSTSARPSOIPTPVNNNTKKRDSKT 2820  
 QY 2821 DSTESSGTQSPKRHSGLVTSV 2843  
 Db 2821 DSTESSGTQSPKRHSGLVTSV 2843  
 RESULT 6  
 US-08-452-655B-2  
 ; Sequence 2, Application US/08452655B  
 ; Patent No. 5783666  
 ; GENERAL INFORMATION:  
 ; APPLICANT: ALBERTSEN, HANS  
 ; APPLICANT: ANAND, RAKESH  
 ; APPLICANT: CARLSON, MARY  
 ; APPLICANT: GRODEN, JOANNA  
 ; APPLICANT: HEDGE, PHILIP J.  
 ; APPLICANT: JOSLYN, GEOFF  
 ; APPLICANT: KINZLER, KENNETH  
 ; APPLICANT: MARKHAM, ALEXANDER F.  
 ; APPLICANT: NAKAMURA, YUSUKE  
 ; APPLICANT: THLIVERIS, ANDREW  
 ; TITLE OF INVENTION: INHERITED AND SOMATIC MUTATIONS OF APC  
 ; TITLE OF INVENTION: GENE IN COLORECTAL CANCER IN HUMANS  
 ; NUMBER OF SEQUENCES: 102  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Banner & Witcoff, Ltd.  
 ; STREET: 1001 G Street, NW  
 ; CITY: Washington  
 ; STATE: D.C.  
 ; COUNTRY: USA  
 ; ZIP: 20001-4598  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/452,655B  
 ; FILING DATE: 25-MAY-1995  
 ; CLASSIFICATION: 530  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 08/289,548  
 ; FILING DATE: 12-AUG-1994  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 07/741,940  
 ; FILING DATE: 08-AUG-1991  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Kagan, Sarah A.  
 ; REGISTRATION NUMBER: 32,141  
 ; REFERENCE/DOCKET NUMBER: 1107.49964  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 202-508-9100  
 ; TELEFAX: 202-508-9299  
 ; INFORMATION FOR SEQ ID NO: 2:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 2843 amino acids  
 ; TYPE: amino acid  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: protein  
 ; US-08-452-655B-2  
 Query Match 99.7%; Score 14533; DB 1; Length 2843;  
 Best Local Similarity 99.8%; Pred. No. 0;  
 Matches 2836; Conservative 2; Mismatches 5; Indels 0; Gaps 0;  
 QY 1 MAAASYDQLLKQVEALKMENSRLRQLEDNSNHLTKLETEASNKMKVQLKQSGTIDEAM 60

Db 1 MAAASYDQLLKQVEALKMENSRLRQLEDNSNHLTKLETEASNKMKVQLKQSGTIDEAM 60  
 QY 61 ASSGQIDLLERLKEINLNDSSNFPKLSKMSLSYSGREGSVSRSGCECPVPMWGSPPR 120  
 Db 61 ASSGQIDLLERLKEINLNDSSNFPKLSKMSLSYSGREGSVSRSGCECPVPMWGSPPR 120  
 QY 121 RGFVNGSRESTGYLBELEKERSLLADLDXBEKDWYIAQLNLTKRIDSPLTENPFL 180  
 Db 121 RGFVNGSRESTGYLBELEKERSLLADLDXBEKDWYIAQLNLTKRIDSPLTENPFL 180  
 QY 181 QDILTRROLEYEARQIRVAMEEQIGTODMEKRAQRRIARIQQIEKDIRLRLQLSQAT 240  
 Db 181 QDILTRROLEYEARQIRVAMEEQIGTODMEKRAQRRIARIQQIEKDIRLRLQLSQAT 240  
 QY 241 EAERSSQNKHETGSHDAERQNEGGVGEINNMATSGNGQGSTTRMDHETASVLSSTHSA 300  
 Db 241 EAERSSQNKHETGSHDAERQNEGGVGEINNMATSGNGQGSTTRMDHETASVLSSTHSA 300  
 QY 301 PRLTSHLGTKEVWYISLLSMLGTHDKDDMSRTLLAMSSQSDSCISMRQSGCLPLLIQLL 360  
 Db 301 PRLTSHLGTKEVWYISLLSMLGTHDKDDMSRTLLAMSSQSDSCISMRQSGCLPLLIQLL 360  
 QY 361 HGNDKDSVLLGNRSGSKEARARASAAALHNIHSQDDKRGRRREIRVLHLEQIRAYCETC 420  
 Db 361 HGNDKDSVLLGNRSGSKEARARASAAALHNIHSQDDKRGRRREIRVLHLEQIRAYCETC 420  
 QY 421 WEQBAHEPGMDQKNPMPAPVEHQICPAVCVLMKLSFDEEHRHAMNELGQLAIAELLQ 480  
 Db 421 WEQBAHEPGMDQKNPMPAPVEHQICPAVCVLMKLSFDEEHRHAMNELGQLAIAELLQ 480  
 QY 481 VDCMYGLTNDHYSITLRRYAGMALTNLTGFDVANKATLCMKGCMRALVAQLKSEEDL 540  
 Db 481 VDCMYGLTNDHYSITLRRYAGMALTNLTGFDVANKATLCMKGCMRALVAQLKSEEDL 540  
 QY 541 QQVIASVLRNLSWRADVNSKTLREVGSVKALMECALEVKESTLKSVALMNLASHCT 600  
 Db 541 QQVIASVLRNLSWRADVNSKTLREVGSVKALMECALEVKESTLKSVALMNLASHCT 600  
 QY 601 ENKADI CAVDGALAPLVTGTLTYRSQNTNTLAIIESGGGILRNVSLLIATNEDHQLRENN 660  
 Db 601 ENKADI CAVDGALAPLVTGTLTYRSQNTNTLAIIESGGGILRNVSLLIATNEDHQLRENN 660  
 QY 661 CLQTLQLKSHSLTIYSNACGLTNLSARNPKDQEAALWDGAVSMKLNLIHSHKHMIAM 720  
 Db 661 CLQTLQLKSHSLTIYSNACGLTNLSARNPKDQEAALWDGAVSMKLNLIHSHKHMIAM 720  
 QY 721 GSAALRNLMANREPAKYKDANIMSPGSSLPVLRKQKALEAELDAQHLSETPDNIOLS 780  
 Db 721 GSAALRNLMANREPAKYKDANIMSPGSSLPVLRKQKALEAELDAQHLSETPDNIOLS 780  
 QY 781 PKASHRSKQRHKQSLGYDVFDTNRHDDNRSDNFTNMTVLPYLNLTVPSSSSSRGS 840  
 Db 781 PKASHRSKQRHKQSLGYDVFDTNRHDDNRSDNFTNMTVLPYLNLTVPSSSSSRGS 840  
 QY 841 LDSSRSKDRSLERERIGLGNVHPATENGTSGKRGLOISTTAAQIAKWEVSAIHTS 900  
 Db 841 LDSSRSKDRSLERERIGLGNVHPATENGTSGKRGLOISTTAAQIAKWEVSAIHTS 900  
 QY 901 QEDRSSGSTTTELHCVTDERNALRRSSAAHTSHNTYNTFKSENSNRITCSMPYAKLEYKRSS 960  
 Db 901 QEDRSSGSTTTELHCVTDERNALRRSSAAHTSHNTYNTFKSENSNRITCSMPYAKLEYKRSS 960  
 QY 961 NDSLNSVSSNDGYGKRCQMKPSIESYSEDEDESCYGYQYPAALAHKIHSANHMDNDGE 1020  
 Db 961 NDSLNSVSSNDGYGKRCQMKPSIESYSEDEDESCYGYQYPAALAHKIHSANHMDNDGE 1020  
 QY 1021 LDTPIYSLKYSDLOLNSGRQSPSONERWARPKHIIDEIKQSEQRQSRNQSTTYPYTE 1080  
 Db 1021 LDTPIYSLKYSDLOLNSGRQSPSONERWARPKHIIDEIKQSEQRQSRNQSTTYPYTE 1080  
 QY 1081 STDDKHLKFPHFQCCQVCPYRSRGANGSETNRVSGNHGINQVNSQLCQEDDDEDDKP 1140

Db 1081 STDDKHLKQPHFGQECVSPVRSRGANGSETNRVGSNHGINQNVQSLCQEDDDYEDDKP 1140  
 QY 1141 TNYSEYSEEEHEEERPTNYSIKYNEEKRHVDQIDYSLKYATDIPSSOKQSFSEFSKS 1200  
 Db 1141 TNYSEYSEEEHEEERPTNYSIKYNEEKRHVDQIDYSLKYATDIPSSOKQSFSEFSKS 1200  
 QY 1201 SGOSSKTEHMSSENSTPSSNAKRONQLHPSSAQSGRGOPOQAATCKYSSINQETIQ 1260  
 Db 1201 SGOSSKTEHMSSENSTPSSNAKRONQLHPSSAQSGRGOPOQAATCKYSSINQETIQ 1260  
 QY 1261 TYCVEDTPICFGRCSLSLSAEDEIGCNOTTOBADSANTLQIAEIKGKIGTRSAEDPV 1320  
 Db 1261 TYCVEDTPICFGRCSLSLSAEDEIGCNOTTOBADSANTLQIAEIKGKIGTRSAEDPV 1320  
 QY 1321 SEVPASVQHPRYKSRRLQSSLSGSARHKAHVEFPSSGAKSPKSGAQTPKSPPEHYVOET 1380  
 Db 1321 SEVPASVQHPRYKSRRLQSSLSGSARHKAHVEFPSSGAKSPKSGAQTPKSPPEHYVOET 1380  
 QY 1381 PLMFGRCTSVSLDFERSIASSVQSPCSQWVGIIISPSDLPSQCTWPPPSRSKTPP 1440  
 Db 1381 PLMFGRCTSVSLDFERSIASSVQSPCSQWVGIIISPSDLPSQCTWPPPSRSKTPP 1440  
 QY 1441 PPQTAQTKREVPKNAKPTAEREGPKQAANAAVQVQLPDADTLLHFATESTPDGF 1500  
 Db 1441 PPQTAQTKREVPKNAKPTAEREGPKQAANAAVQVQLPDADTLLHFATESTPDGF 1500  
 QY 1501 SCSSLSALSDBPFIQKDVLRIMPPVOENDNGNETSEQPKESNENQKEAEKTIIDSE 1560  
 Db 1501 SCSSLSALSDBPFIQKDVLRIMPPVOENDNGNETSEQPKESNENQKEAEKTIIDSE 1560  
 QY 1561 KOLLDDSDDDDEIIEECIIISAMPTKSSRKKKPAQTASKLPPPVARPKSPQVLYKLLPS 1620  
 Db 1561 KOLLDDSDDDDEIIEECIIISAMPTKSSRKKKPAQTASKLPPPVARPKSPQVLYKLLPS 1620  
 QY 1621 QNRLOPKHVSPTDGDMPRVYCVGPTPINFSTATSLDLTIESPNELAAEGVGRGAQ 1680  
 Db 1621 QNRLOPKHVSPTDGDMPRVYCVGPTPINFSTATSLDLTIESPNELAAEGVGRGAQ 1680  
 QY 1681 SEFEKEDTIPTEGRSTDEAQGGKTSVTPILDDNKAEBEGDILAEACINSAMPKCKSHK 1740  
 Db 1681 SEFEKEDTIPTEGRSTDEAQGGKTSVTPILDDNKAEBEGDILAEACINSAMPKCKSHK 1740  
 QY 1741 FRVKIMDOVQOASASSAPNKNQLDGKKKTSVPKPIQNTYRTRVRKNKADSKNNLN 1800  
 Db 1741 FRVKIMDOVQOASASSAPNKNQLDGKKKTSVPKPIQNTYRTRVRKNKADSKNNLN 1800  
 QY 1801 ABRVSDNKKDKKQNLKNNKDFNDKLPNNEDRVGSAFDPSPHYTPIEGTPYCFSEND 1860  
 Db 1801 ABRVSDNKKDKKQNLKNNKDFNDKLPNNEDRVGSAFDPSPHYTPIEGTPYCFSEND 1860  
 QY 1861 SLSSLDFFDDDDVDLSREKAEALRKAENKESAEKVTSTELTNSQOSANKTQAIKQPINR 1920  
 Db 1861 SLSSLDFFDDDDVDLSREKAEALRKAENKESAEKVTSTELTNSQOSANKTQAIKQPINR 1920  
 QY 1921 GQPKPILOKQSTFPOSSKDIIPRGATDEKLOFAIENTPVCFSHNSLSLSLSDIDQENN 1980  
 Db 1921 GQPKPILOKQSTFPOSSKDIIPRGATDEKLOFAIENTPVCFSHNSLSLSLSDIDQENN 1980  
 QY 1981 NKENEPIKETEPPDSOGEPKQASGYAPKSPHVEDTPVCFGRNSLSLSLSDSDDDLLQ 2040  
 Db 1981 NKENEPIKETEPPDSOGEPKQASGYAPKSPHVEDTPVCFGRNSLSLSLSDSDDDLLQ 2040  
 QY 2041 ECISAMPKKKPSRLKGNEXKHSPPNMGGIIGEDLTLDKDIQRPDSHGLSPDSENF 2100  
 Db 2041 ECISAMPKKKPSRLKGNEXKHSPPNMGGIIGEDLTLDKDIQRPDSHGLSPDSENF 2100  
 QY 2101 WKAIQEGANSIVSSLHQAAAACLSRQASDSDSILSLKSGISLGSPPHLPDQEKPT 2160  
 Db 2101 WKAIQEGANSIVSSLHQAAAACLSRQASDSDSILSLKSGISLGSPPHLPDQEKPT 2160  
 QY 2161 SNKGPRILKPGKSTLETKKIESESGIKGKKVYKSLITGKVRNSIISGQMKOPLQAN 2220  
 Db 2161 SNKGPRILKPGKSTLETKKIESESGIKGKKVYKSLITGKVRNSIISGQMKOPLQAN 2220

QY 2221 MPISIRGRWTMIHIPGVNRNSSSTSPVSKKGPPLKTPASKSPSEGTATTSPRGAKPSVKS 2280  
 Db 2221 MPISIRGRWTMIHIPGVNRNSSSTSPVSKKGPPLKTPASKSPSEGTATTSPRGAKPSVKS 2280  
 QY 2281 ELSPVARQTSOIGSSKAPSRSGRSTPSRPAQPLSRPIQSPGRNSISPGRNGISPPN 2340  
 Db 2281 ELSPVARQTSOIGSSKAPSRSGRSTPSRPAQPLSRPIQSPGRNSISPGRNGISPPN 2340  
 QY 2341 KLSQLPTTSPSTASTKSSGSGKMSYTSPOKQSQQLTKOTGLSKNASSIPRSEASKG 2400  
 Db 2341 KLSQLPTTSPSTASTKSSGSGKMSYTSPOKQSQQLTKOTGLSKNASSIPRSEASKG 2400  
 QY 2401 LQNMNNGANKKVELSRMSSTKSSSESSESRPVLVROSTFIKEAPSTLRKLEESA 2460  
 Db 2401 LQNMNNGANKKVELSRMSSTKSSSESSESRPVLVROSTFIKEAPSTLRKLEESA 2460  
 QY 2461 SFESLSPPSPASPTRSOAQTPLVLSPLPDMWLSSTHSSVQAGGWRKLPPLNLSPTIEYNDG 2520  
 Db 2461 SFESLSPPSPASPTRSOAQTPLVLSPLPDMWLSSTHSSVQAGGWRKLPPLNLSPTIEYNDG 2520  
 QY 2521 RPAKRHDIAARSHESPSRLPINRSGTWKRHSHSSSLPRVSTWRTTGGSSSILSASSES 2580  
 Db 2521 RPAKRHDIAARSHESPSRLPINRSGTWKRHSHSSSLPRVSTWRTTGGSSSILSASSES 2580  
 QY 2581 SEKAKSEDEKHVNSISGTKOSKENOVSAKGTWKIKENEFSPNTSQTVSSGATNGAES 2640  
 Db 2581 SEKAKSEDEKHVNSISGTKOSKENOVSAKGTWKIKENEFSPNTSQTVSSGATNGAES 2640  
 QY 2641 KTLIYQMAPAVSKTDDVWRIEDCPINNPRSGSPGTNTPPVIDSVSEKANPNIKDSKDN 2700  
 Db 2641 KTLIYQMAPAVSKTDDVWRIEDCPINNPRSGSPGTNTPPVIDSVSEKANPNIKDSKDN 2700  
 QY 2701 QAKONVNGSVPMRTVGLNRLTSFOVDAPDQKTEIKPGQNNPVVSTNESPVERT 2760  
 Db 2701 QAKONVNGSVPMRTVGLNRLTSFOVDAPDQKTEIKPGQNNPVVSTNESPVERT 2760  
 QY 2761 PFSSSSSSKHSPPSGTVAARVTPFNPNPSRKSADTSARPQIPTPVNNNTKKRDSKT 2820  
 Db 2761 PFSSSSSSKHSPPSGTVAARVTPFNPNPSRKSADTSARPQIPTPVNNNTKKRDSKT 2820  
 QY 2821 DSTESSGTQPKRHSGSYLTVS 2843  
 Db 2821 DSTESSGTQPKRHSGSYLTVS 2843

RESULT 7

; Sequence 7, Application US/08452655B  
 ; Patent No. 5783666  
 ; GENERAL INFORMATION:

; APPLICANT: ALBERTSEN, HANS  
 ; APPLICANT: ANAND, RAKESH  
 ; APPLICANT: CARLSON, MARY  
 ; APPLICANT: GRODEN, JOANNA  
 ; APPLICANT: HEDGE, PHILIP J.  
 ; APPLICANT: JOSLYN, GEOFF  
 ; APPLICANT: KINZLER, KENNETH  
 ; APPLICANT: MARKHAM, ALEXANDER F.  
 ; APPLICANT: NAKAMURA, YUSUKE  
 ; APPLICANT: THLIVERIS, ANDREW  
 ; TITLE OF INVENTION: INHERITED AND SOMATIC MUTATIONS OF APC  
 ; TITLE OF INVENTION: GENE IN COLORECTAL CANCER IN HUMANS  
 ; NUMBER OF SEQUENCES: 102  
 ; CORRESPONDENCE ADDRESS:  
 ; STREET: 1001 G Street, NW  
 ; CITY: Washington  
 ; STATE: D.C.  
 ; COUNTRY: USA  
 ; ZIP: 20001-4598  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk



```

; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/452,655B
; FILING DATE: 25-MAY-1995
; CLASSIFICATION: 530
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: US 08/289,548
; FILING DATE: 12-AUG-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/741,940
; FILING DATE: 08-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Kagan, Sarah A.
; REGISTRATION NUMBER: 32,141
; REFERENCE/DOCKET NUMBER: 1107.49964
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-508-9100
; TELEFAX: 202-508-9299
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2843 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEICAL: YES
; ANTI-SENSE: NO
; US-08-452-655B-7

```

Query Match 99.7%; Score 14533; DB 1; Length 2843;  
 Best Local Similarity 99.8%; Pred. No. 0;  
 Matches 2836; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

```

QY 1 MAASYPQLLKQVEALWENSLRQELNLSNHLTKLETSASNMKEVLKQCSIDEAM 60
DB 1 MAASYPQLLKQVEALWENSLRQELNLSNHLTKLETSASNMKEVLKQCSIDEAM 60
QY 61 ASSQIDILLERKELNLDSSNFPQVGLRSKMSLSYSGRSGSVSSRSGECPVPMGSFPR 120
DB 61 ASSQIDILLERKELNLDSSNFPQVGLRSKMSLSYSGRSGSVSSRSGECPVPMGSFPR 120
QY 121 RGFVNGRSTGYLEELKERSLLADLDEEKEKDWYAAQLQNLTKRIDSLPLTENFSL 180
DB 121 RGFVNGRSTGYLEELKERSLLADLDEEKEKDWYAAQLQNLTKRIDSLPLTENFSL 180
QY 181 QTDLTRQLEYEARQIRVAMEEQLGTQDMKEKRAORRIARIQOIEKDILRIQLLOSQAT 240
DB 181 QTDLTRQLEYEARQIRVAMEEQLGTQDMKEKRAORRIARIQOIEKDILRIQLLOSQAT 240
QY 241 EAERSSQNKHETGSHDAERQEGGVGEINATSGNGCGSTTRMDHETASVLSSSTHSA 300
DB 241 EAERSSQNKHETGSHDAERQEGGVGEINATSGNGCGSTTRMDHETASVLSSSTHSA 300
QY 301 PRLTSLHGTKEVMYSLMLGTHDKDDMSRTLAMSSQDSCINRQSGCPLLIQLL 360
DB 301 PRLTSLHGTKEVMYSLMLGTHDKDDMSRTLAMSSQDSCINRQSGCPLLIQLL 360
QY 361 HGNDKDSVLLGNRSGSKEARASAAALHNIHQDPDKRREIRVLHLLEQIRAYCETC 420
DB 361 HGNDKDSVLLGNRSGSKEARASAAALHNIHQDPDKRREIRVLHLLEQIRAYCETC 420
QY 421 WEQEAEPGMDQKNPMPAVEHQICPACVLMKLSFDEHRHAMNGLGGLQAIALLQ 480
DB 421 WEQEAEPGMDQKNPMPAVEHQICPACVLMKLSFDEHRHAMNGLGGLQAIALLQ 480
QY 481 VDCEMYGLTNDHYSTILRRVAGMALTNLTGFDVANKATLCSMKGCMREALVAQLKSESDL 540
DB 481 VDCEMYGLTNDHYSTILRRVAGMALTNLTGFDVANKATLCSMKGCMREALVAQLKSESDL 540
QY 541 QOVIASVLRNLSWRADVNSKKTLEREVSVKALMECALEVKKESLKSVALNLSAHT 600

```

```

DB 541 QOVIASVLRNLSWRADVNSKKTLEREVSVKALMECALEVKKESLKSVALNLSAHT 600
QY 601 ENKADICAVDGLAPLVGTLVRSQNTNLAIESGGILRVNSSLIATNEHROILRENN 660
DB 601 ENKADICAVDGLAPLVGTLVRSQNTNLAIESGGILRVNSSLIATNEHROILRENN 660
QY 661 CLQTLQHLKSHSLTIVSNAGCTLWLSARNPQKQALWMDMGAVSMLKNLHSHKHMIAM 720
DB 661 CLQTLQHLKSHSLTIVSNAGCTLWLSARNPQKQALWMDMGAVSMLKNLHSHKHMIAM 720
QY 721 GSAAALRNLMANRPKAKYKDANIMSPGSLPSLHVRSKQKALFAELDAQHLSFTFNIDLS 780
DB 721 GSAAALRNLMANRPKAKYKDANIMSPGSLPSLHVRSKQKALFAELDAQHLSFTFNIDLS 780
QY 781 PKASHRSKORHKQSLYGDYVFDTRHDDNRSDNNTGNMTVLSPYLNTTVLPSSSSRGS 840
DB 781 PKASHRSKORHKQSLYGDYVFDTRHDDNRSDNNTGNMTVLSPYLNTTVLPSSSSRGS 840
QY 841 LDSSRSSEKDRSLERERIGLGNTHPATENPCTSSKRGQLQISTTAAQIAKVMEEVSAHTS 900
DB 841 LDSSRSSEKDRSLERERIGLGNTHPATENPCTSSKRGQLQISTTAAQIAKVMEEVSAHTS 900
QY 901 QEDRSSGTTTELHCVTDERNALRRSSAAHTSNNTYNTKSENSNRTCSMPYAKLEYKRSS 960
DB 901 QEDRSSGTTTELHCVTDERNALRRSSAAHTSNNTYNTKSENSNRTCSMPYAKLEYKRSS 960
QY 961 NDSLNSVSNDCYKRGQMKPSIESYSEDDSKFCYQYQADLAHKIHSANHMDNDGE 1020
DB 961 NDSLNSVSNDCYKRGQMKPSIESYSEDDSKFCYQYQADLAHKIHSANHMDNDGE 1020
QY 1021 LDTPINYSKYSDEQLNSGRSPSONERWARPKHIIIEDEIKOSQOROSNSTYTPVYTE 1080
DB 1021 LDTPINYSKYSDEQLNSGRSPSONERWARPKHIIIEDEIKOSQOROSNSTYTPVYTE 1080
QY 1081 STDDXHLKAFPHFGQECVSPYRSGANGSETNVSNGHGINQVNSLQCEDDYEDDKP 1140
DB 1081 STDDXHLKAFPHFGQECVSPYRSGANGSETNVSNGHGINQVNSLQCEDDYEDDKP 1140
QY 1141 TNSYRYSEEEQHEEERPTNYSIKYNEEKRVDPIDYSLKYATDIPSSOKQSFSSKS 1200
DB 1141 TNSYRYSEEEQHEEERPTNYSIKYNEEKRVDPIDYSLKYATDIPSSOKQSFSSKS 1200
QY 1201 SSGQSKTEHWSSENSTPSSNAKQNLHPSSAQSRSQCPKAAATCKVSSINQETIQ 1260
DB 1201 SSGQSKTEHWSSENSTPSSNAKQNLHPSSAQSRSQCPKAAATCKVSSINQETIQ 1260
QY 1261 TYCVEDTPICFSRCSLSLSSAEDEICNCTTQEADSANTLQIAEIKGIGTSAEDPV 1320
DB 1261 TYCVEDTPICFSRCSLSLSSAEDEICNCTTQEADSANTLQIAEIKGIGTSAEDPV 1320
QY 1321 SEVPASVQHPRTKSRLOGSSLSSESARHKAVERPSSGAKSPKSGAOTPKSPHYVQET 1380
DB 1321 SEVPASVQHPRTKSRLOGSSLSSESARHKAVERPSSGAKSPKSGAOTPKSPHYVQET 1380
QY 1381 PLMFRCSTSVSLDSFESRSIASSVQSEPCSGMVGIIISPSDLDPSCQOTMPPSRSTPP 1440
DB 1381 PLMFRCSTSVSLDSFESRSIASSVQSEPCSGMVGIIISPSDLDPSCQOTMPPSRSTPP 1440
QY 1441 PPQTAQTKREVKNKAPTASRESGPKQAANVAQVQVLPADTLHLHFATSTPDGF 1500
DB 1441 PPQTAQTKREVKNKAPTASRESGPKQAANVAQVQVLPADTLHLHFATSTPDGF 1500
QY 1501 SCSSLSLSALSDDEPIQKDELIRMPVQENDNGNETESEOPEKSENQEKAEKTIIDSE 1560
DB 1501 SCSSLSLSALSDDEPIQKDELIRMPVQENDNGNETESEOPEKSENQEKAEKTIIDSE 1560
QY 1561 KOLLDDSDDDIIELEECIIISAMPTKSRKQKPAQATASKLPPPVARKPSOLPYKLLPS 1620
DB 1561 KOLLDDSDDDIIELEECIIISAMPTKSRKQKPAQATASKLPPPVARKPSOLPYKLLPS 1620
QY 1621 QNRLOPQKHVSFTPGDDMPRVYCVEGTPINFSTATSLDITIESPPNELAAGEVGRGAQ 1680
DB 1621 QNRLOPQKHVSFTPGDDMPRVYCVEGTPINFSTATSLDITIESPPNELAAGEVGRGAQ 1680

```



Qy 1681 SGEFKRDTIPIEGRSTDEAOGKTSVYTIPELDNKAEEGDILAEBCINSAMPKCKSHK 1740  
Db 1681 SGEFKRDTIPIEGRSTDEAOGKTSVYTIPELDNKAEEGDILAEBCINSAMPKCKSHK 1740  
Qy 1741 FRVKIMDQVQASASSAPNKNQLDGKKKPTSPVKPIPQNTYRTVRKNADSKNNLN 1800  
Db 1741 FRVKIMDQVQASASSAPNKNQLDGKKKPTSPVKPIPQNTYRTVRKNADSKNNLN 1800  
Qy 1801 AERFSDNKKQNLKNNKDPNDKLPNNEDRVGSAFDPSPHYTPIEGTYPYCFSRND 1860  
Db 1801 AERFSDNKKQNLKNNKDPNDKLPNNEDRVGSAFDPSPHYTPIEGTYPYCFSRND 1860  
Qy 1861 SLSSLDFFDDDDVLDLREKAEKAKENKESAKVTSHLTNSQOSANKTOAIKQPINR 1920  
Db 1861 SLSSLDFFDDDDVLDLREKAEKAKENKESAKVTSHLTNSQOSANKTOAIKQPINR 1920  
Qy 1921 GQPKILOKQSTFPOSSKDIIPRGAADEKQNFPAIENTPVCFSNSSLSDIDQNNN 1980  
Db 1921 GQPKILOKQSTFPOSSKDIIPRGAADEKQNFPAIENTPVCFSNSSLSDIDQNNN 1980  
Qy 1981 NKNEPIKETEPDPSQGGPKQASGYAPKSPHVEDTPVCFSRNSSLSDIDSEDDLLQ 2040  
Db 1981 NKNEPIKETEPDPSQGGPKQASGYAPKSPHVEDTPVCFSRNSSLSDIDSEDDLLQ 2040  
Qy 2041 ECISAMPKKKPSRLKGDNEKHSRNMGGILGEDTLTLDKDIQRPDSEHGLSPDSENF 2100  
Db 2041 ECISAMPKKKPSRLKGDNEKHSRNMGGILGEDTLTLDKDIQRPDSEHGLSPDSENF 2100  
Qy 2101 WKAIQEGANSIVSSLHQAACLRQASSDSILSLKSGISLGSPEHLPDQOEKPT 2160  
Db 2101 WKAIQEGANSIVSSLHQAACLRQASSDSILSLKSGISLGSPEHLPDQOEKPT 2160  
Qy 2161 SNKGRILKPKGKSTLETKKIBESKGIKGGKVKYKSLITGKVRNSNISQGMQKQLOAN 2220  
Db 2161 SNKGRILKPKGKSTLETKKIBESKGIKGGKVKYKSLITGKVRNSNISQGMQKQLOAN 2220  
Qy 2221 MPSISGRMTIHIPGVRNSSSTSPVSKGPPKTPASKSPSEGGTATTPRGAKPSVKS 2280  
Db 2221 MPSISGRMTIHIPGVRNSSSTSPVSKGPPKTPASKSPSEGGTATTPRGAKPSVKS 2280  
Qy 2281 ELSPVAROTSLQGGSKAPSRSGSDSTPSRPAQPLSRPIQSPGRNISIPGRNGISPPN 2340  
Db 2281 ELSPVAROTSLQGGSKAPSRSGSDSTPSRPAQPLSRPIQSPGRNISIPGRNGISPPN 2340  
Qy 2341 KLSQLPRTSSPTASTKSGSGKMYTSPGRQMSQONLTQTLGSKNASSIPRSESASKG 2400  
Db 2341 KLSQLPRTSSPTASTKSGSGKMYTSPGRQMSQONLTQTLGSKNASSIPRSESASKG 2400  
Qy 2401 LNMNNGCANKKVELSRMSTKSGSDRSERPVLRQSTFIKEAPSPILRRKLEESA 2460  
Db 2401 LNMNNGCANKKVELSRMSTKSGSDRSERPVLRQSTFIKEAPSPILRRKLEESA 2460  
Qy 2461 SPESLSPSSRPASPTRSQAQTPVLSPLPDMSLTHSSVQAGGWRKLPNLSPTIEYNDG 2520  
Db 2461 SPESLSPSSRPASPTRSQAQTPVLSPLPDMSLTHSSVQAGGWRKLPNLSPTIEYNDG 2520  
Qy 2521 RPAKHDIARSHSESPRLPINRGTWKRHSKSSSLPRVSTWRTGSSSSSILSASSES 2580  
Db 2521 RPAKHDIARSHSESPRLPINRGTWKRHSKSSSLPRVSTWRTGSSSSSILSASSES 2580  
Qy 2581 SEKAKSEDEKHVNSISGTQKSENOVSAGTWKIKENEFPSTNSTQTVSSGATNGAES 2640  
Db 2581 SEKAKSEDEKHVNSISGTQKSENOVSAGTWKIKENEFPSTNSTQTVSSGATNGAES 2640  
Qy 2641 KTLIYQMAPAVSKTEDVWVRIEDCPINNPRSGRSTGNTPPVIDSVSEKAPNPKDSKDN 2700  
Db 2641 KTLIYQMAPAVSKTEDVWVRIEDCPINNPRSGRSTGNTPPVIDSVSEKAPNPKDSKDN 2700  
Qy 2701 QAKQNVGSGVPMRTVGLENRLTSLFIQVADPQKTEIKPGONNPVPSVETNESPIVERT 2760  
Db 2701 QAKQNVGSGVPMRTVGLENRLTSLFIQVADPQKTEIKPGONNPVPSVETNESPIVERT 2760

Qy 2761 PFSSSSSKHSSPSGTVAARVTPFNPNPSPRKSSADTSARPSQIPTPVNNNTKKRDSKT 2820  
Db 2761 PFSSSSSKHSSPSGTVAARVTPFNPNPSPRKSSADTSARPSQIPTPVNNNTKKRDSKT 2820  
Qy 2821 DSTESSGTQSPKHSGLSYLTVS 2843  
Db 2821 DSTESSGTQSPKHSGLSYLTVS 2843

RESULT 8  
US-08-450-582-2  
; Sequence 2, Application US/08450582  
; Patent No. 6114124  
; GENERAL INFORMATION:  
; APPLICANT: ALBERTSEN, HANS  
; APPLICANT: ANAND, RAKESH  
; APPLICANT: CARLSON, MARY  
; APPLICANT: GRODEN, JOANNA  
; APPLICANT: HEDGE, PHILIP J.  
; APPLICANT: JOSLYN, GEOFF  
; APPLICANT: KINZLER, KENNETH  
; APPLICANT: MARKHAM, ALEXANDER F.  
; APPLICANT: NAKAMURA, YUSUKE  
; APPLICANT: THLIVERIS, ANDREW  
; TITLE OF INVENTION: INHERITED AND SOMATIC MUTATIONS OF APC  
; TITLE OF INVENTION: GENE IN COLORECTAL CANCER IN HUMANS  
; NUMBER OF SEQUENCES: 102  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Banner & Witcoff, Ltd.  
; STREET: 1001 G Street, NW  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20001-4598  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/450,582  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/452,655  
; FILING DATE: 25-MAY-1995  
; APPLICATION NUMBER: US 08/289,548  
; FILING DATE: 12-AUG-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/741,940  
; FILING DATE: 08-AUG-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Kagan, Sarah A.  
; REGISTRATION NUMBER: 32,141  
; REFERENCE/DOCKET NUMBER: 1107.49964  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-508-9100  
; TELEFAX: 202-508-9239  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2843 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-450-582-2

Query Match 99.7%; Score 14533; DB 3; Length 2843;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 2836; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Qy 1 MAAASYDOLLKQVLEALKNWENSLRQELDENSEHNLTKLETSANMKVLEKOLQGSIEDAM 60  
Db 1 MAAASYDOLLKQVLEALKNWENSLRQELDENSEHNLTKLETSANMKVLEKOLQGSIEDAM 60

QY 61 ASSQIDLLERLKEINLNDNFPGVKLRKMSLRSGREGSVSSRGSGCSPVPMGSPFR 120  
DB 61 ASSQIDLLERLKEINLNDNFPGVKLRKMSLRSGREGSVSSRGSGCSPVPMGSPFR 120  
QY 121 RGFGVSGRESTGYLEBELEKRSLLADLXKEEKEKWYQAQLNHLKRIIDSLPLTENFSL 180  
DB 121 RGFGVSGRESTGYLEBELEKRSLLADLXKEEKEKWYQAQLNHLKRIIDSLPLTENFSL 180  
QY 181 QTLTRQLEYEARQIRVAMEEOLGTQDMKXKAQRRIARIQOIEKIDILRIROLLSQAT 240  
DB 181 QTMTRQLEYEARQIRVAMEEOLGTQDMKXKAQRRIARIQOIEKIDILRIROLLSQAT 240  
QY 241 EABRSSQKHETGSHDAERONEGGVGEINMATSGNGQGGTTRMDHETASVLSSSSTHSA 300  
DB 241 EABRSSQKHETGSHDAERONEGGVGEINMATSGNGQGGTTRMDHETASVLSSSSTHSA 300  
QY 301 PRRLTSHLGTKEVMYSLLSMLGTHDXDMSRTLILAMSSQDSCISMROGCLPLLIQLL 360  
DB 301 PRRLTSHLGTKEVMYSLLSMLGTHDXDMSRTLILAMSSQDSCISMROGCLPLLIQLL 360  
QY 361 HGNDKDSVLLGNSRGSKEARASAAALHNIHSPQDDKRGRRRIRVLHLEQIRAYCETC 420  
DB 361 HGNDKDSVLLGNSRGSKEARASAAALHNIHSPQDDKRGRRRIRVLHLEQIRAYCETC 420  
QY 421 WEQEAHEPGMDOKNMPAPVEHQICPAVCVLMKLSFDEEHRHAMELGGLOIAIELLO 480  
DB 421 WEQEAHEPGMDOKNMPAPVEHQICPAVCVLMKLSFDEEHRHAMELGGLOIAIELLO 480  
QY 481 VDCMYGLTNDHYGITTIRRYAGMALTNLTFDGVANKATLCMSKGMCFALVAQLKSEEDL 540  
DB 481 VDCMYGLTNDHYGITTIRRYAGMALTNLTFDGVANKATLCMSKGMCFALVAQLKSEEDL 540  
QY 541 QOVTASVLRNLWSRADVNSKTLREVGSVKALMECALEVKESTLKSVALNLSAHC 600  
DB 541 QOVTASVLRNLWSRADVNSKTLREVGSVKALMECALEVKESTLKSVALNLSAHC 600  
QY 601 ENKADI CAVDGALAFVGLTYRSQTNLTALIESGGILRNVSLLATNEDHRRQILRENN 660  
DB 601 ENKADI CAVDGALAFVGLTYRSQTNLTALIESGGILRNVSLLATNEDHRRQILRENN 660  
QY 661 CLQTLLOHLKSHSUTIVSNACGTLNLSARNPKDQEAALMDMGAVSMKNLIHSHKHMIAM 720  
DB 661 CLQTLLOHLKSHSUTIVSNACGTLNLSARNPKDQEAALMDMGAVSMKNLIHSHKHMIAM 720  
QY 721 GSAAALRNLMANRPKYKDNINISPGSLSPLHVRKOKALEAEADACHLSETFDNIDNLS 780  
DB 721 GSAAALRNLMANRPKYKDNINISPGSLSPLHVRKOKALEAEADACHLSETFDNIDNLS 780  
QY 781 PKASHRSKQRHQSGLYGDYVFTNRHDDNRSDFNTGNMTVLSPLYLNTTVLPSSSSRGS 840  
DB 781 PKASHRSKQRHQSGLYGDYVFTNRHDDNRSDFNTGNMTVLSPLYLNTTVLPSSSSRGS 840  
QY 841 LDSRSEKORSLERGIGLGNHPATENPCTSKRGLOISTTAAQIAKVMESVAIHTS 900  
DB 841 LDSRSEKORSLERGIGLGNHPATENPCTSKRGLOISTTAAQIAKVMESVAIHTS 900  
QY 901 QEDRSSGTTTELHCVTDERNALRSSAAHTSHNTYNTFKSENSNRCTCSMPYAKLEYKRS 960  
DB 901 QEDRSSGTTTELHCVTDERNALRSSAAHTSHNTYNTFKSENSNRCTCSMPYAKLEYKRS 960  
QY 961 NDSLNSVSSNDGYKRCQMKPESIESYEDDESXFCYQYPADLAHKIHSANHMDNDGE 1020  
DB 961 NDSLNSVSSNDGYKRCQMKPESIESYEDDESXFCYQYPADLAHKIHSANHMDNDGE 1020  
QY 1021 LDTPIYNSLKYSDQLNSGRQSPQNERWARPKHIIIDEIKQSEQRQSRNQSTTYPVYTE 1080  
DB 1021 LDTPIYNSLKYSDQLNSGRQSPQNERWARPKHIIIDEIKQSEQRQSRNQSTTYPVYTE 1080  
QY 1081 STDDXHLKFPQHFQOQCVSPYRSRGRANGSETNRVGSNHGINONVQSGLCQEDDYEDDXP 1140  
DB 1081 STDDXHLKFPQHFQOQCVSPYRSRGRANGSETNRVGSNHGINONVQSGLCQEDDYEDDXP 1140

QY 1141 TNSERYSEEBQHEEERPTNYSIKYNEKRVHVDQPIDYSLKYATDIPSSQKQSFSSKS 1200  
DB 1141 TNSERYSEEBQHEEERPTNYSIKYNEKRVHVDQPIDYSLKYATDIPSSQKQSFSSKS 1200  
QY 1201 SSGQSKTEHWSSESSENTSTPSSNAKRONLHPSAOSRSGOPOKAATCKVSSINQETIQ 1260  
DB 1201 SSGQSKTEHWSSESSENTSTPSSNAKRONLHPSAOSRSGOPOKAATCKVSSINQETIQ 1260  
QY 1261 TYCVEDTPICFRSRCSLSLSAEDICNQTTOEADSANTLQIAEIKGTGTRSAEDPV 1320  
DB 1261 TYCVEDTPICFRSRCSLSLSAEDICNQTTOEADSANTLQIAEIKGTGTRSAEDPV 1320  
QY 1321 SEVPASVQHPRTKSRLOGSSLSSESARHKAVERPSGAKSPSKSGAOTPKSPHYVQET 1380  
DB 1321 SEVPASVQHPRTKSRLOGSSLSSESARHKAVERPSGAKSPSKSGAOTPKSPHYVQET 1380  
QY 1381 PLMFSRCTSVSLSLDSFESRSIASSVQSEPCSGMVSGIISPSDLPSDQTMPPSRSKTTP 1440  
DB 1381 PLMFSRCTSVSLSLDSFESRSIASSVQSEPCSGMVSGIISPSDLPSDQTMPPSRSKTTP 1440  
QY 1441 PPQTAQTKREVPPKNAKPTABKRESGPKQAANAARVQVLPDADTLLHPATSTPDGF 1500  
DB 1441 PPQTAQTKREVPPKNAKPTABKRESGPKQAANAARVQVLPDADTLLHPATSTPDGF 1500  
QY 1501 SCSSLSLSALSDDEPFIQKDVLRIMPVQENDNGNETESEQPKESNENQEKAEKTIIDSE 1560  
DB 1501 SCSSLSLSALSDDEPFIQKDVLRIMPVQENDNGNETESEQPKESNENQEKAEKTIIDSE 1560  
QY 1561 KOLLDDSDDDDIIELEECIIISAMPTKSRKOKPAQATASKLPPPVARKPSOLPVYKLLPS 1620  
DB 1561 KOLLDDSDDDDIIELEECIIISAMPTKSRKOKPAQATASKLPPPVARKPSOLPVYKLLPS 1620  
QY 1621 QNRLOPQKHVSFTGDDMPRVYCVEGTPINFSTATSLDITIESPPELAAAGEVRGGAQ 1680  
DB 1621 QNRLOPQKHVSFTGDDMPRVYCVEGTPINFSTATSLDITIESPPELAAAGEVRGGAQ 1680  
QY 1681 SGEFEKRTDITTEGRSTDEAOGGKTSVTIPELDDNKAEEGDIIAECINSAMPKGSHPK 1740  
DB 1681 SGEFEKRTDITTEGRSTDEAOGGKTSVTIPELDDNKAEEGDIIAECINSAMPKGSHPK 1740  
QY 1741 FRVKIMDOVQOASASSAPNKOIDGKKKXPTSPVKPIPONTYRTRVRKADSKNNLN 1800  
DB 1741 FRVKIMDOVQOASASSAPNKOIDGKKKXPTSPVKPIPONTYRTRVRKADSKNNLN 1800  
QY 1801 AERFSDNKKSKQNLKNNKDFNDKLPNNEDRVGSGFAFDSPHHYTPIEGTPYCFSRND 1860  
DB 1801 AERFSDNKKSKQNLKNNKDFNDKLPNNEDRVGSGFAFDSPHHYTPIEGTPYCFSRND 1860  
QY 1861 SLSLDFDDDVDVLSREKAEIRKAKENKESPAKTSHTELTSNOQSANKTOAIKQPINR 1920  
DB 1861 SLSLDFDDDVDVLSREKAEIRKAKENKESPAKTSHTELTSNOQSANKTOAIKQPINR 1920  
QY 1921 GQPKFILOKQSTFPQSSKDI PDRGAATDEKLQNFIAENTPVCFSHNSLSLSLSDIDQENN 1980  
DB 1921 GQPKFILOKQSTFPQSSKDI PDRGAATDEKLQNFIAENTPVCFSHNSLSLSLSDIDQENN 1980  
QY 1981 NKNEPIKETEPPOSGQSPKQASGVAPKSFHVEDTPVCFSRNSSLSLSIDSEDDLQ 2040  
DB 1981 NKNEPIKETEPPOSGQSPKQASGVAPKSFHVEDTPVCFSRNSSLSLSIDSEDDLQ 2040  
QY 2041 ECISAMPKKKKSRLLKGDNEKHSFRNMGGLGEBDLTLDLKDIOKDPSEHGLSPDSNFD 2100  
DB 2041 ECISAMPKKKKSRLLKGDNEKHSFRNMGGLGEBDLTLDLKDIOKDPSEHGLSPDSNFD 2100  
QY 2101 WKAIQEGANSIVSSILHQAAAAACLSRQASSDSILSKSGISLGSFPFHTPDQEEKPFT 2160  
DB 2101 WKAIQEGANSIVSSILHQAAAAACLSRQASSDSILSKSGISLGSFPFHTPDQEEKPFT 2160  
QY 2161 SNKGRIILKPEKSTLETFTKTESKGIKGGKVKYKSLITGKVRNSNEISGQMKQPIQAN 2220  
DB 2161 SNKGRIILKPEKSTLETFTKTESKGIKGGKVKYKSLITGKVRNSNEISGQMKQPIQAN 2220  
QY 2221 MFSIRSGRTMIHIFGVRNRSSTSVSKGKGPPLKTPASKSPSEGOATTSPRGAKPSVKS 2280

Db 2221 MFISISGRTHIHPGVRNSSSTSPVSKGPKLPKTPASKSPSEGTATTSPRGAKPSVKS 2280  
Qy 2281 ELSVAROTSQIGGSKAPSRGSRDSTPSRPAQOPLSRPIQSPGRNIGISPOGNGISPPN 2340  
Db 2281 ELSVAROTSQIGGSKAPSRGSRDSTPSRPAQOPLSRPIQSPGRNIGISPOGNGISPPN 2340  
Qy 2341 KLSQLPRTSSPSTASTKSSGSKMYSTSPCRMSQOQLTKQGLSKNASSIPRSESASKG 2400  
Db 2341 KLSQLPRTSSPSTASTKSSGSKMYSTSPCRMSQOQLTKQGLSKNASSIPRSESASKG 2400  
Qy 2401 LNMQNGGANKKVELSRMSSSTKSSGSDRSERPVLRQSTFIKEAPSPILRRKLEESA 2460  
Db 2401 LNMQNGGANKKVELSRMSSSTKSSGSDRSERPVLRQSTFIKEAPSPILRRKLEESA 2460  
Qy 2461 SFESLSPSRPASPTRSQOATPVLSPLPDMSLSTHSSVQAGGWRKLPNLSPTIYNDG 2520  
Db 2461 SFESLSPSRPASPTRSQOATPVLSPLPDMSLSTHSSVQAGGWRKLPNLSPTIYNDG 2520  
Qy 2521 RPAKHDIARSHSESPRLPINRGSTWKRHSHSSSLPRVSTWRTGSSSILSASSES 2580  
Db 2521 RPAKHDIARSHSESPRLPINRGSTWKRHSHSSSLPRVSTWRTGSSSILSASSES 2580  
Qy 2581 SEKAKSEDEKHVNSISGTHQSKENQVSAKGTWRKIKENEFSPNTSTQTVSSGATNGAES 2640  
Db 2581 SEKAKSEDEKHVNSISGTHQSKENQVSAKGTWRKIKENEFSPNTSTQTVSSGATNGAES 2640  
Qy 2641 KTLIYQMAPAVSKTEDEVVRIEDCPINNPRSGSPGTNTPVIDSVSEKAPNKKSDN 2700  
Db 2641 KTLIYQMAPAVSKTEDEVVRIEDCPINNPRSGSPGTNTPVIDSVSEKAPNKKSDN 2700  
Qy 2701 QAKQNVGSGVPMRTVGLNRLTSTFIQVDAPQKTEIKPGQNNPVPVSETNESPIVERT 2760  
Db 2701 QAKQNVGSGVPMRTVGLNRLTSTFIQVDAPQKTEIKPGQNNPVPVSETNESPIVERT 2760  
Qy 2761 PFSSSSSKHSPSGTVAARVTPFNYNPSPRKSSADSTSARPSQIPTPVNNTKGRDST 2820  
Db 2761 PFSSSSSKHSPSGTVAARVTPFNYNPSPRKSSADSTSARPSQIPTPVNNTKGRDST 2820  
Qy 2821 DSTESSGTSQPKRHSGSYLVTSV 2843  
Db 2821 DSTESSGTSQPKRHSGSYLVTSV 2843

RESULT 9

US-08-450-582-7

; Sequence 7, Application US/08450582.

; Patent No. 6114124

; GENERAL INFORMATION:

; APPLICANT: ALBERTSEN, HANS

; APPLICANT: ANAND, RAKESH

; APPLICANT: CARLSON, MARY

; APPLICANT: GRODEN, JOANNA

; APPLICANT: HEDGE, PHILIP J.

; APPLICANT: JOSLYN, GEOFF

; APPLICANT: KINZLER, KENNETH

; APPLICANT: MARKHAM, ALEXANDER F.

; APPLICANT: NAKAMURA, YUSUKE

; APPLICANT: THLIVIERIS, ANDREW

; TITLE OF INVENTION: INHERITED AND SOMATIC MUTATIONS OF APC

; TITLE OF INVENTION: GENE IN COLORECTAL CANCER IN HUMANS

; NUMBER OF SEQUENCES: 102

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Banner & Witcoff, Ltd.

; STREET: 1001 G Street, NW

; CITY: Washington

; STATE: D.C.

; COUNTRY: USA

; ZIP: 20001-4598

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; FILING DATE: US/08/450,582

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/452,655

; FILING DATE: 25-MAY-1995

; APPLICATION NUMBER: US 08/289,548

; FILING DATE: 12-AUG-1994

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/741,940

; FILING DATE: 08-AUG-1991

; ATTORNEY/AGENT INFORMATION:

; NAME: kagan, Sarah A.

; REGISTRATION NUMBER: 32,141

; REFERENCE/DOCKET NUMBER: 1107.49964

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 202-508-9100

; TELEFAX: 202-508-9299

; INFORMATION FOR SEQ ID NO: 7:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 2843 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; HYPOTHEetical: YES

; ANTI-SENSE: NO

; US-08-450-582-7

Query Match 99.7%; Score 14533; DB 3; Length 2843;

Best Local Similarity 99.8%; Pred. No. 0;

Matches 2836; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Qy 1 MAAASYDQLLKQVEALKWENSNLROLEDNSNHLTKLETEASNKVEVLKOLQSIIDEAM 60  
Db 1 MAAASYDQLLKQVEALKWENSNLROLEDNSNHLTKLETEASNKVEVLKOLQSIIDEAM 60  
Qy 61 ASSGOIDLLRLKELNLDSSNFPQVKLRKMSLRSGVSGREGSVSSSGEGSPVPMGSFPR 120  
Db 61 ASSGOIDLLRLKELNLDSSNFPQVKLRKMSLRSGVSGREGSVSSSGEGSPVPMGSFPR 120  
Qy 121 RGFVNGSRESTGYLELEKERSILLADLDKEEKEKWYQAOLQNLTKRIDLPLTENFSL 180  
Db 121 RGFVNGSRESTGYLELEKERSILLADLDKEEKEKWYQAOLQNLTKRIDLPLTENFSL 180  
Qy 181 QTDLTREOLEYEAQIRVAVEEQLGTCODMEKRAQRIARIQIIEKDIRIOLLOSCAT 240  
Db 181 QTDLTREOLEYEAQIRVAVEEQLGTCODMEKRAQRIARIQIIEKDIRIOLLOSCAT 240  
Qy 241 EAERSQNKHETGSHDAERQNEGGVGEINMATSGNGQGGSTTRMDHETASVLSSTHSA 300  
Db 241 EAERSQNKHETGSHDAERQNEGGVGEINMATSGNGQGGSTTRMDHETASVLSSTHSA 300  
Qy 301 PRLLTSHLGTKEVWYVSLLSMLGTHDKDDMSRILLAMSSQSDSCISMRQSGCLPLLIQLL 360  
Db 301 PRLLTSHLGTKEVWYVSLLSMLGTHDKDDMSRILLAMSSQSDSCISMRQSGCLPLLIQLL 360  
Qy 361 HGNDKDSVLLGNSRGSKAERARASAAHLNIIHSQPDQKRGRRIRVILHLEQIRAYCETC 420  
Db 361 HGNDKDSVLLGNSRGSKAERARASAAHLNIIHSQPDQKRGRRIRVILHLEQIRAYCETC 420  
Qy 421 WEMQEAHEPGMDQKQMPAPVSHQICPAVCVLMKLSFDEEHRHAMNELGGLQAIALLQ 480  
Db 421 WEMQEAHEPGMDQKQMPAPVSHQICPAVCVLMKLSFDEEHRHAMNELGGLQAIALLQ 480  
Qy 481 VDCEMYGLTNDHYSITLRRYAGMALTNLTFDGVANKATLCSMKGCKRALVAQKSSIDL 540  
Db 481 VDCEMYGLTNDHYSITLRRYAGMALTNLTFDGVANKATLCSMKGCKRALVAQKSSIDL 540  
Qy 541 QQVIASVLRNLSWRADVNSKKTIREVGSVFALMECALEVKESTLKSVALNLSAHT 600

Db 541 QOVIASVLRNLSWRADYNSKTLREVGSVKALMECALEVKESTLKSIVLSALNLSAHTC 600  
 Qy 601 ENKADICAVDICALAFVLTITYSQTNLTALIESGGILRNVSLLIATNEDHROIARENN 660  
 Db 601 ENKADICAVDICALAFVLTITYSQTNLTALIESGGILRNVSLLIATNEDHROIARENN 660  
 Qy 661 CLOTLLQHLKSHSLTIVSNACGTLWNLARNPKDOEALWDMGAVSMLKNLIHSHKHMIAM 720  
 Db 661 CLOTLLQHLKSHSLTIVSNACGTLWNLARNPKDOEALWDMGAVSMLKNLIHSHKHMIAM 720  
 Qy 721 GSAALRNLMANRPAYKNDANIMSPGSLPSLHVKKQKALEAELDAQHSETFDNIDNLS 780  
 Db 721 GSAALRNLMANRPAYKNDANIMSPGSLPSLHVKKQKALEAELDAQHSETFDNIDNLS 780  
 Qy 781 PKASHRSKQKHQSLYGDYVFDNRHDDNRSDFNFTGNMTVLSFYLNTTVLPSSSSSRGS 840  
 Db 781 PKASHRSKQKHQSLYGDYVFDNRHDDNRSDFNFTGNMTVLSFYLNTTVLPSSSSSRGS 840  
 Qy 841 LDSSRSKORSLRERGIGLGNVHPATENPCTGSKGLQISTTAAOIAKYMEEVSAIHTS 900  
 Db 841 LDSSRSKORSLRERGIGLGNVHPATENPCTGSKGLQISTTAAOIAKYMEEVSAIHTS 900  
 Qy 901 QEDRSSGTTTELHCVTDERNALRRSSAAHSTHNTYFTKSENRTCSMPYAKLEYKRSS 960  
 Db 901 QEDRSSGTTTELHCVTDERNALRRSSAAHSTHNTYFTKSENRTCSMPYAKLEYKRSS 960  
 Qy 961 NDSLNSVSSNDGYKRGOMKPSIESYEDDESFCGYQPADLAHKSANHMDNDGE 1020  
 Db 961 NDSLNSVSSNDGYKRGOMKPSIESYEDDESFCGYQPADLAHKSANHMDNDGE 1020  
 Qy 1021 LDTPIYLSKYDQSLNAGRSQSPONERWARPKHIIIEDEIKQSEORSRNQSTTYPVYTE 1080  
 Db 1021 LDTPIYLSKYDQSLNAGRSQSPONERWARPKHIIIEDEIKQSEORSRNQSTTYPVYTE 1080  
 Qy 1081 STDDKHLKQFHFQOQECVSPYRSRGANGSETNRVGSNHGINOVQSGLCQEDDYEDDKP 1140  
 Db 1081 STDDKHLKQFHFQOQECVSPYRSRGANGSETNRVGSNHGINOVQSGLCQEDDYEDDKP 1140  
 Qy 1141 TNYSERVSEBQHEBERPTNYSIKYNEEKHVDQPIDYSLKYATDIPSSQKQSFPSKS 1200  
 Db 1141 TNYSERVSEBQHEBERPTNYSIKYNEEKHVDQPIDYSLKYATDIPSSQKQSFPSKS 1200  
 Qy 1201 SSGQSKTEHWSSESSTTSPSSNAKQNLHPSSAQSRSGQPKAATCKVSSINQETIQ 1260  
 Db 1201 SSGQSKTEHWSSESSTTSPSSNAKQNLHPSSAQSRSGQPKAATCKVSSINQETIQ 1260  
 Qy 1261 TYCVEDTPICFSRCSLSLSAABDEIGCQNTQOADSANTLOIAETIKGIGTRESADPV 1320  
 Db 1261 TYCVEDTPICFSRCSLSLSAABDEIGCQNTQOADSANTLOIAETIKGIGTRESADPV 1320  
 Qy 1321 SEVPAVQHPRTKSSRLQSSLSSESARHKAVERFPSSGAKSPSKGAQTPKSPPEHYVQET 1380  
 Db 1321 SEVPAVQHPRTKSSRLQSSLSSESARHKAVERFPSSGAKSPSKGAQTPKSPPEHYVQET 1380  
 Qy 1381 PLMFSRCTSVSSLDSFESRSIASSVQSEPCSGMWGIIISPSDLPDSGOTMPPSRSKTTP 1440  
 Db 1381 PLMFSRCTSVSSLDSFESRSIASSVQSEPCSGMWGIIISPSDLPDSGOTMPPSRSKTTP 1440  
 Qy 1441 PPPQTAQTKREVPKNKAFTAEKRESGPKQAANAAVQOVLPDADTLHLFAESTPDGF 1500  
 Db 1441 PPPQTAQTKREVPKNKAFTAEKRESGPKQAANAAVQOVLPDADTLHLFAESTPDGF 1500  
 Qy 1501 SCSSLSALSLEDPFIQKVDELRLMPVQNDNGNETESEPQKESNENQKEAEKTI DSE 1560  
 Db 1501 SCSSLSALSLEDPFIQKVDELRLMPVQNDNGNETESEPQKESNENQKEAEKTI DSE 1560  
 Qy 1561 KDLDDSDDDDDIEILEBCCIISAMPTKSRKGPQAOTASKLPPPVARKPSQLPVYKLLPS 1620  
 Db 1561 KDLDDSDDDDDIEILEBCCIISAMPTKSRKGPQAOTASKLPPPVARKPSQLPVYKLLPS 1620  
 Qy 1621 QNRLOPQKHVSFTPGDDMPRVYCVEGTPIINFSTATSLSDLTIESPPNELAAGSVRGGAQ 1680  
 Db 1621 QNRLOPQKHVSFTPGDDMPRVYCVEGTPIINFSTATSLSDLTIESPPNELAAGSVRGGAQ 1680

Qy 1681 SGEFEKRTDITPEGRSTDEAOGGKTSSVTIPELDDNKAEBGDILAEICINSAMPKGSHPK 1740  
 Db 1681 SGEFEKRTDITPEGRSTDEAOGGKTSSVTIPELDDNKAEBGDILAEICINSAMPKGSHPK 1740  
 Qy 1741 FRVKKINDOVQOASASSAPNKNOLDGKKKPTSPVKPIPONTREYRTRVRKNADSKNNLN 1800  
 Db 1741 FRVKKINDOVQOASASSAPNKNOLDGKKKPTSPVKPIPONTREYRTRVRKNADSKNNLN 1800  
 Qy 1801 AERVFSDNKSCKKQNLKNNSKDFNDKLPNNEDRVGSFAFDSPHHYPIEGTYPYCFSRND 1860  
 Db 1801 AERVFSDNKSCKKQNLKNNSKDFNDKLPNNEDRVGSFAFDSPHHYPIEGTYPYCFSRND 1860  
 Qy 1861 SLSLDFDDDDVDLSREKAEIRKAKENKESFAKYTSHTELTSNQSANKTQAKAQPINR 1920  
 Db 1861 SLSLDFDDDDVDLSREKAEIRKAKENKESFAKYTSHTELTSNQSANKTQAKAQPINR 1920  
 Qy 1921 CQPKPILQKQSTFPQSSKDIIPDRGAATDEKLQNFPAIENTPVCFSHNSLSLSDIDQENN 1980  
 Db 1921 CQPKPILQKQSTFPQSSKDIIPDRGAATDEKLQNFPAIENTPVCFSHNSLSLSDIDQENN 1980  
 Qy 1981 NKEPEPIKETFPDSQGEPPKQASGYAPKSFHVEDTVPFCSRNSLSLSLSDIDQENN 2040  
 Db 1981 NKEPEPIKETFPDSQGEPPKQASGYAPKSFHVEDTVPFCSRNSLSLSLSDIDQENN 2040  
 Qy 2041 ECISAMPKKKPRLKGDNEKHSFRNNGGILGEBDLTLDLKDIOQRPDSEHGLSPDSENF 2100  
 Db 2041 ECISAMPKKKPRLKGDNEKHSFRNNGGILGEBDLTLDLKDIOQRPDSEHGLSPDSENF 2100  
 Qy 2101 WKAIQEGANSIVSLHQAAAAACLSROASSDSISILSKGISLGSFPHLTPDOEKPF 2160  
 Db 2101 WKAIQEGANSIVSLHQAAAAACLSROASSDSISILSKGISLGSFPHLTPDOEKPF 2160  
 Qy 2161 SNKGPRILLKPEKSTLETKTIESKGIKGGKVVYKSLITGKVSNSIEISQMKQPIQAN 2220  
 Db 2161 SNKGPRILLKPEKSTLETKTIESKGIKGGKVVYKSLITGKVSNSIEISQMKQPIQAN 2220  
 Qy 2221 MPISIRGRMTIHPGVNRSSSSTSPVSKGPPKLPKASPSECOATTPRGAAPSVKS 2280  
 Db 2221 MPISIRGRMTIHPGVNRSSSSTSPVSKGPPKLPKASPSECOATTPRGAAPSVKS 2280  
 Qy 2281 ELSVPARTSOIGSSKAPSRSGRSDTTPRPAQOPLSRPIQSGRNSISPRNGISPPN 2340  
 Db 2281 ELSVPARTSOIGSSKAPSRSGRSDTTPRPAQOPLSRPIQSGRNSISPRNGISPPN 2340  
 Qy 2341 KLSQLPRTSSPSTASTKSSGSKNSYTPSGRQMSQNLTKQTGLSKNASSIPRSESASKG 2400  
 Db 2341 KLSQLPRTSSPSTASTKSSGSKNSYTPSGRQMSQNLTKQTGLSKNASSIPRSESASKG 2400  
 Qy 2401 LQNMNGANKKVELSRMSSTKSSGSDSRSEPRVLRQSTPFKEAPSPTRRLKRESA 2460  
 Db 2401 LQNMNGANKKVELSRMSSTKSSGSDSRSEPRVLRQSTPFKEAPSPTRRLKRESA 2460  
 Qy 2461 SPESLSRSPASPTRSQAQTPVLSPLSDMSLSTHSSVQAGGWRKLPNPLSPRIEYNDG 2520  
 Db 2461 SPESLSRSPASPTRSQAQTPVLSPLSDMSLSTHSSVQAGGWRKLPNPLSPRIEYNDG 2520  
 Qy 2521 RPAKHDIARSHSPSPRLPINRSGTWKHSKHSSSLPRVSTWRTGSSSSILSASSES 2580  
 Db 2521 RPAKHDIARSHSPSPRLPINRSGTWKHSKHSSSLPRVSTWRTGSSSSILSASSES 2580  
 Qy 2581 SEKAKSEDEKHNSTSGTKQSKENQVSAKGTWRKI KENEPSTNSTQTVSSGATNGAES 2640  
 Db 2581 SEKAKSEDEKHNSTSGTKQSKENQVSAKGTWRKI KENEPSTNSTQTVSSGATNGAES 2640  
 Qy 2641 KTLIYQMAPAVSKTEDVWVRIEDCPINNPRSGRSGPTGNTTPTVIDSVSEKANPNTKSKDN 2700  
 Db 2641 KTLIYQMAPAVSKTEDVWVRIEDCPINNPRSGRSGPTGNTTPTVIDSVSEKANPNTKSKDN 2700  
 Qy 2701 QAKQNVGVSGVPMRTVGLNENLTSFIOVADPQKGTETKPGQNNPVPVSETNESPIVERT 2760  
 Db 2701 QAKQNVGVSGVPMRTVGLNENLTSFIOVADPQKGTETKPGQNNPVPVSETNESPIVERT 2760

QY 781 PKASHRSQRKHQSLYGYVFDTRHDDNRSDNFRNTGNTVLSPLYLNTTVLPSSSSSRGS 840  
DB 781 PKASHRSQRKHQSLYGYVFDTRHDDNRSDNFRNTGNTVLSPLYLNTTVLPSSSSSRGS 840  
QY 841 LDSRSEKDRSLERBERGICLGNYPATENPGTSSKRGLOISTTAAQIAKWEEVSAIHTS 900  
DB 841 LDSRSEKDRSLERBERGICLGNYPATENPGTSSKRGLOISTTAAQIAKWEEVSAIHTS 900  
QY 901 QEDRSSGGTTLHCVTDERNALRSSAAHTSNFTYNTFKTSNSNRTCSMPYAKLEYKRSS 960  
DB 901 QEDRSSGGTTLHCVTDERNALRSSAAHTSNFTYNTFKTSNSNRTCSMPYAKLEYKRSS 960  
QY 961 NDSLNSVSSSGYKRGQWKPSIESYSEDDSKSCSYGOYPADIAHKIHSANHMDNDGE 1020  
DB 961 NDSLNSVSSSGYKRGQWKPSIESYSEDDSKSCSYGOYPADIAHKIHSANHMDNDGE 1020  
QY 1021 LDTPIYSLKYSDQLNSGRQSPSQRERWARPKHIIIEIKQSEQRQSRNQSTTYPVYTE 1080  
DB 1021 LDTPIYSLKYSDQLNSGRQSPSQRERWARPKHIIIEIKQSEQRQSRNQSTTYPVYTE 1080  
QY 1081 STDDKHLKFPQHFQGOECVPSYRSGANGSETNRVGSNHGINQVSOGLCOEDDYEDDKP 1140  
DB 1081 STDDKHLKFPQHFQGOECVPSYRSGANGSETNRVGSNHGINQVSOGLCOEDDYEDDKP 1140  
QY 1141 TNSERYSEEEQHEBERPTNYSIKYNEEKRHVDOPIDYSLKYATDIPSSQKQSFSSKS 1200  
DB 1141 TNSERYSEEEQHEBERPTNYSIKYNEEKRHVDOPIDYSLKYATDIPSSQKQSFSSKS 1200  
QY 1201 SSGQSKTEHMSSESENTSTPSSNAKRONQLHPSSAQSRSGOPQKAATCKVSSINQETIQ 1260  
DB 1201 SSGQSKTEHMSSESENTSTPSSNAKRONQLHPSSAQSRSGOPQKAATCKVSSINQETIQ 1260  
QY 1261 TYCVEDTPICFRCSLSLSAEBEIGCNQTTQEADSANTLOIAEIKEKIGTSAEDPV 1320  
DB 1261 TYCVEDTPICFRCSLSLSAEBEIGCNQTTQEADSANTLOIAEIKEKIGTSAEDPV 1320  
QY 1321 SEVPAVSHPRKTSRLQCGSLSSSARHKAVEFSSGAKSPSKGAQTPKPPPEHYVOET 1380  
DB 1321 SEVPAVSHPRKTSRLQCGSLSSSARHKAVEFSSGAKSPSKGAQTPKPPPEHYVOET 1380  
QY 1381 PLMFRCSTSVSLDSFERSSTASSVQSEPCSGMVSGIISPDLPSPCGTQTPPERSKTPP 1440  
DB 1381 PLMFRCSTSVSLDSFERSSTASSVQSEPCSGMVSGIISPDLPSPCGTQTPPERSKTPP 1440  
QY 1441 PPQTAQTKREVPKNAKPTAEKRESGPKQAAVNAAVQVQVLPDADTLLHFATESTPDGF 1500  
DB 1441 PPQTAQTKREVPKNAKPTAEKRESGPKQAAVNAAVQVQVLPDADTLLHFATESTPDGF 1500  
QY 1501 SCSSLSLSALSDPEPIQKDVLRIMPVQENDNGNETESBOPKESNENQEKAEKTIIDSE 1560  
DB 1501 SCSSLSLSALSDPEPIQKDVLRIMPVQENDNGNETESBOPKESNENQEKAEKTIIDSE 1560  
QY 1561 KDLLDDSDDDDIIEILEECIIISAMPTKSRKAKKPAQTASKLPPPVARXPSOLPVYKLLPS 1620  
DB 1561 KDLLDDSDDDDIIEILEECIIISAMPTKSRKAKKPAQTASKLPPPVARXPSOLPVYKLLPS 1620  
QY 1621 QNRLOPQKHVSFTPGDDMPRVYCVEGTPINFSTATSLDSTLIESPPNELAAGEVGRGAQ 1680  
DB 1621 QNRLOPQKHVSFTPGDDMPRVYCVEGTPINFSTATSLDSTLIESPPNELAAGEVGRGAQ 1680  
QY 1681 SCEPEKRTDPIPTGSRSTDAQGGKTSVTVIPELDDNKAEEGDILAEACINSAMPKGSHPK 1740  
DB 1681 SCEPEKRTDPIPTGSRSTDAQGGKTSVTVIPELDDNKAEEGDILAEACINSAMPKGSHPK 1740  
QY 1741 FRVKIMDQVQASASSAPNKNQLDGKKKFTSPVKPIQNTYRTRVRKNADSKNNLN 1800  
DB 1741 FRVKIMDQVQASASSAPNKNQLDGKKKFTSPVKPIQNTYRTRVRKNADSKNNLN 1800  
QY 1801 AERVFSNDKSKQNLKNNSKDFNDKLPNNEDRVGRSFAFDPSPHYTPIEGTFCFSRND 1860  
DB 1801 AERVFSNDKSKQNLKNNSKDFNDKLPNNEDRVGRSFAFDPSPHYTPIEGTFCFSRND 1860

RESULT 8

AAR63508

ID AAR63508

XX

standard; protein; 2842 AA.

QY 1861 SLSSLDFFDDDDVLSPEKAELRKAKENKSEAKVTSHTELTSTNOQSANKTQAIKAPINR 1920  
DB 1861 SLSSLDFFDDDDVLSPEKAELRKAKENKSEAKVTSHTELTSTNOQSANKTQAIKAPINR 1920  
QY 1921 GQFKPILQKQSTPPQSSKDIPIRGAATDEKLNFAIENTPVCFSHNSLSLSDIDQENN 1980  
DB 1921 GQFKPILQKQSTPPQSSKDIPIRGAATDEKLNFAIENTPVCFSHNSLSLSDIDQENN 1980  
QY 1981 NKENBPIKETEPPDSDGEBPSKQASGYAPKSHVEDTIVCFPSRNSLSLSLSDIDQENN 2040  
DB 1981 NKENBPIKETEPPDSDGEBPSKQASGYAPKSHVEDTIVCFPSRNSLSLSLSDIDQENN 2040  
QY 2041 ECTSSAMPKPKKPSRLKGDNEKHSPRNMGILGEDLTLDLKOIQRDPSHGLSPDSENF 2100  
DB 2041 ECTSSAMPKPKKPSRLKGDNEKHSPRNMGILGEDLTLDLKOIQRDPSHGLSPDSENF 2100  
QY 2101 WKAIQEGANSIVSSLHQAAAAACLSRQASDSDSILSLKSGISLGSFPHLTPDQEEKPFT 2160  
DB 2101 WKAIQEGANSIVSSLHQAAAAACLSRQASDSDSILSLKSGISLGSFPHLTPDQEEKPFT 2160  
QY 2161 SNKGPRILKPGEKSTLETTKIESKIGGKVKYKSLITGKVRNSSEISGQWKQLOAN 2220  
DB 2161 SNKGPRILKPGEKSTLETTKIESKIGGKVKYKSLITGKVRNSSEISGQWKQLOAN 2220  
QY 2221 MPSISGRGTMIHIPGVNRSSTSPVSKGPPPLKTPASKSPSEGTATTPSPRGAKPSVK 2280  
DB 2221 MPSISGRGTMIHIPGVNRSSTSPVSKGPPPLKTPASKSPSEGTATTPSPRGAKPSVK 2280  
QY 2281 ELSFVARQTSIQIGSKAPSRSGRSDTSPRPAQQLSRPIQSPGRNSISPGRNGISPPN 2340  
DB 2281 ELSFVARQTSIQIGSKAPSRSGRSDTSPRPAQQLSRPIQSPGRNSISPGRNGISPPN 2340  
QY 2341 KLSQLPRTSPSTASTKSGSGRMSYTSFGRQVSOQNLTKQTGLSKNASSIPRSEASKG 2400  
DB 2341 KLSQLPRTSPSTASTKSGSGRMSYTSFGRQVSOQNLTKQTGLSKNASSIPRSEASKG 2400  
QY 2401 LNOMNNGANKVKVELSRMSTKSSSESSEDRSRPVLVRQSTFIKAPSTLRKLEESA 2460  
DB 2401 LNOMNNGANKVKVELSRMSTKSSSESSEDRSRPVLVRQSTFIKAPSTLRKLEESA 2460  
QY 2461 SFESLSPSSRPASPTRSQAQTPVLSPLDMSLSTHSSVQAGGWRKLPNLSPTIEYNDG 2520  
DB 2461 SFESLSPSSRPASPTRSQAQTPVLSPLDMSLSTHSSVQAGGWRKLPNLSPTIEYNDG 2520  
QY 2521 RPAKRHDIAARSHSESPRLPINRSGTWKREHSHSSLPVSTWRTTSGSSSILSASSES 2580  
DB 2521 RPAKRHDIAARSHSESPRLPINRSGTWKREHSHSSLPVSTWRTTSGSSSILSASSES 2580  
QY 2581 SEKAKSDEKHVNSISGTQKSKENQVSAKTWKIKENEFSPNTNSTQTVSSGATNGAES 2640  
DB 2581 SEKAKSDEKHVNSISGTQKSKENQVSAKTWKIKENEFSPNTNSTQTVSSGATNGAES 2640  
QY 2641 KTLIYQMAPAVSKTEDVWVRIEDCPINNPRSGRSPGTGNTPPVIDSVSEKANPNIKDSKON 2700  
DB 2641 KTLIYQMAPAVSKTEDVWVRIEDCPINNPRSGRSPGTGNTPPVIDSVSEKANPNIKDSKON 2700  
QY 2701 QAKQNVGNGVPMRTVGLNENLNSFTQVDAPDQKGTIEIKPGQNNPVPVSETNESSIVERT 2760  
DB 2701 QAKQNVGNGVPMRTVGLNENLNSFTQVDAPDQKGTIEIKPGQNNPVPVSETNESSIVERT 2760  
QY 2761 PFSSSSSSKHSXSSPGTVAARVTFNNPNPSRKSADSTARPQOIPTPVNNNTKKRDSKT 2820  
DB 2761 PFSSSSSSKHSXSSPGTVAARVTFNNPNPSRKSADSTARPQOIPTPVNNNTKKRDSKT 2820  
QY 2821 DSTESSGTOQPKHSGSYLVTSV 2843  
DB 2821 DSTESSGTOQPKHSGSYLVTSV 2843

AC AAR63508;  
 XX 25-MAR-2003 (revised)  
 DT 23-MAY-1995 (first entry)  
 XX  
 DE Adenomatous polyposis coli tumour repressor.  
 XX  
 KW Adenomatous polyposis coli; tumour repressor; Gardner's syndrome;  
 KW familial adenomatous polyposis; cancer diagnosis and prognosis;  
 KW tumorigenesis suppression.  
 XX  
 OS Homo sapiens.  
 XX  
 PN US5352775-A.  
 XX  
 PD 04-OCT-1994.  
 XX  
 PF 08-AUG-1991; 91US-00741940.  
 XX  
 PR 16-JAN-1991; 91GB-00000962.  
 PR 16-JAN-1991; 91GB-00000963.  
 PR 16-JAN-1991; 91GB-00000974.  
 PR 16-JAN-1991; 91GB-00000975.  
 XX  
 XX (ICIL ) IMPERIAL CHEM IND.  
 PA (CANC-) CANCER INST.  
 PA (UJJO ) UNIV JOHNS HOPKINS.  
 PA (UTAH ) UNIV UTAH.  
 XX  
 PI Markham AF, Hedge PJ, Anand R, Nakamura Y, Groden J, Kinzler K;  
 PI Thliveris A, Carlson M, Vogelstein B, Albertsen H, White RL;  
 PI Joslyn G;  
 XX  
 DR WPI; 1994-316233/39.  
 DR N-PSDB; AAQ72297.  
 XX  
 PT New human adenomatous polyposis coli DNA encoding tumour repressor - and  
 PT derived primers and probes for diagnosis, prognosis and treatment of  
 PT cancer.  
 XX  
 PS Claim 3; Col 71-84; 113pp; English.  
 XX  
 CC AAQ72297 is a cDNA isolated from the human adenomatous polyposis coli  
 CC (APC) gene, it encodes the tumour repressors described in AAR63507 and  
 CC AAR63508. Determination of alterations in APC or its expression products,  
 CC can be used for the diagnosis and prognosis of cancer e.g. colorectal,  
 CC lung and breast tumours; and for determining predisposition to certain  
 CC cancers such as familial adenomatous polyposis (FAP) and Gardner's  
 CC syndrome. The wild type APC gene (or a part of it) can be used  
 CC therapeutically to restore gene function, while primers and probes  
 CC derived from the cDNA (AAQ72333-400 and AAQ72541-568) can be used to  
 CC detect mutations. Also APC proteins or analogues can be administered to  
 CC compensate for a defective gene, and epithelial cells, or transgenic  
 CC animals carrying a mutated APC allele are useful for detecting  
 CC therapeutic agents able to suppress tumorigenesis. (Updated on 25-MAR-  
 CC 2003 to correct PA field.)  
 XX  
 SQ Sequence 2842 AA;  
 Query Match 99.9%; Score 14548.5; DB 2; Length 2842;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 2842; Conservative 0; Mismatches 0; Indels 1; Gaps 1;  
 QY 1 MAAASDQLLKQVEALQWENSNLRQLEEDNSNLTTLKLETEASNMKVLKQLOGSIDEAM 60  
 DB 1 MAAASDQLLKQVEALQWENSNLRQLEEDNSNLTTLKLETEASNMKVLKQLOGSIDEAM 60  
 QY 61 ASSGQIDLLERLKLNLDSNFFGVKLRKQWLSRYSVSGREGSVSSRGCSVPVWGSFPR 120  
 DB 61 ASSGQIDLLERLKLNLDSNFFGVKLRKQWLSRYSVSGREGSVSSRGCSVPVWGSFPR 120  
 QY 121 RGFVNGSRESTGYLEBELEKERSLLADLDKEEKEKDWYQAQLQNLTKRIDSL-LTENFSL 179

DB 121 RGFVNGSRESTGYLEBELEKERSLLADLDKEEKEKDWYQAQLQNLTKRIDSL-LTENFSL 179  
 QY 181 QTDMTTRQLEYEARQIRVAMEEQLTQCDMEKRAORRIARIQIEKDIILRIQLQSQAT 240  
 DB 180 QTDMTTRQLEYEARQIRVAMEEQLTQCDMEKRAORRIARIQIEKDIILRIQLQSQAT 239  
 QY 241 EAERSSQNKHETGSHDAERQNEGQVGEINWATSGNGGSGTTRMDHETASVLSSSSTHSA 300  
 DB 240 EAERSSQNKHETGSHDAERQNEGQVGEINWATSGNGGSGTTRMDHETASVLSSSSTHSA 299  
 QY 301 PRRLTSHLGTKEVMYVYILLSMLGTHDKDMSRTLLAMSSSQDSCISMRSQSGCLPLLIQLL 360  
 DB 300 PRRLTSHLGTKEVMYVYILLSMLGTHDKDMSRTLLAMSSSQDSCISMRSQSGCLPLLIQLL 359  
 QY 361 HGNDKDSVLLGNSRGSKEARARASAAALHNIHSOPDDKRGRRREIRVHLHLEQIRAYCETC 420  
 DB 360 HGNDKDSVLLGNSRGSKEARARASAAALHNIHSOPDDKRGRRREIRVHLHLEQIRAYCETC 419  
 QY 421 WEQEAHEPGMDQDNMPAPVEHQICPAVCVLMKLSFDEEHRHANNELGGLQAIABLLQ 480  
 DB 420 WEQEAHEPGMDQDNMPAPVEHQICPAVCVLMKLSFDEEHRHANNELGGLQAIABLLQ 479  
 QY 481 VDCMYGLTNDHYSITLRRYAGMALTNLTFCGVANKATLCMSKGCMPALVAQKSESEDL 540  
 DB 480 VDCMYGLTNDHYSITLRRYAGMALTNLTFCGVANKATLCMSKGCMPALVAQKSESEDL 539  
 QY 541 QOVIASVLRNLSWRADVNSKTLREVGSVKALMECALEVKESTLKSVLASALNLSAHT 600  
 DB 540 QOVIASVLRNLSWRADVNSKTLREVGSVKALMECALEVKESTLKSVLASALNLSAHT 599  
 QY 601 ENKADICAVDQALAFVLTGTYRSQTNLTALIESGGGILRVNSLIATNEDHRQILRENN 660  
 DB 600 ENKADICAVDQALAFVLTGTYRSQTNLTALIESGGGILRVNSLIATNEDHRQILRENN 659  
 QY 661 CLQTLLOHLKSHSLTVSNACGLTNLSARNPKDOEALMDGAVSMKLNLIHSGHKMIAM 720  
 DB 660 CLQTLLOHLKSHSLTVSNACGLTNLSARNPKDOEALMDGAVSMKLNLIHSGHKMIAM 719  
 QY 721 GSAALRNLMANRPAPKDKANIMSPGSSLPSPHVRKQKALEAEIDAQHLSETFDNDLS 780  
 DB 720 GSAALRNLMANRPAPKDKANIMSPGSSLPSPHVRKQKALEAEIDAQHLSETFDNDLS 779  
 QY 781 PKASHRSKQHKQSLYGDYVDTNRHDDNRPNTGNTVLPYLNTTVLPSSSSRGS 840  
 DB 780 PKASHRSKQHKQSLYGDYVDTNRHDDNRPNTGNTVLPYLNTTVLPSSSSRGS 839  
 QY 841 LDSRSKDRSLERERGITGLGNVHPATENPCTSSKRGLOISTTAAQIAKVEEVSIAHTS 900  
 DB 840 LDSRSKDRSLERERGITGLGNVHPATENPCTSSKRGLOISTTAAQIAKVEEVSIAHTS 899  
 QY 901 QEDRSSGTTTELHCVTDERNALRRSSAAHNTSYNTFKSNGSNRTCSMPYAKLEYKRSS 960  
 DB 900 QEDRSSGTTTELHCVTDERNALRRSSAAHNTSYNTFKSNGSNRTCSMPYAKLEYKRSS 959  
 QY 961 NDSLSNVSSSDGYGKRGOMKPSIESYSEDDSEKFCPSYGOYPADLAHKIHSANHMDNDGE 1020  
 DB 960 NDSLSNVSSSDGYGKRGOMKPSIESYSEDDSEKFCPSYGOYPADLAHKIHSANHMDNDGE 1019  
 QY 1021 LDTPIYSLKYSDQLNSGRQSPQNERWARPKHIIIEDEIKQSEQRQSRNQSTPYVTE 1080  
 DB 1020 LDTPIYSLKYSDQLNSGRQSPQNERWARPKHIIIEDEIKQSEQRQSRNQSTPYVTE 1079  
 QY 1081 STDDKHLKFPQHQEQECVPSYRSGANGSETNRVGSNHGINQVNSQLCQEDDYEDDKP 1140  
 DB 1080 STDDKHLKFPQHQEQECVPSYRSGANGSETNRVGSNHGINQVNSQLCQEDDYEDDKP 1139  
 QY 1141 TNYSERYSSEEEHEEERPTNYSIKYNEEKRVHDQPIDYSLKYATDIPSSQKSFSPKS 1200  
 DB 1140 TNYSERYSSEEEHEEERPTNYSIKYNEEKRVHDQPIDYSLKYATDIPSSQKSFSPKS 1199  
 QY 1201 SSGOSSKTEHMSSSSENTSTPSSNAKQNLHPSSAQSRSGQPKATCKVSSINQETIQ 1260  
 DB 1200 SSGOSSKTEHMSSSSENTSTPSSNAKQNLHPSSAQSRSGQPKATCKVSSINQETIQ 1259

QY 1261 TYCVEDTTPICRSKSSLSAABEIGCNQTTOTBADANTLQIAEIKEKIGTSAEDPV 1320  
Db 1260 TYCVEDTTPICRSKSSLSAABEIGCNQTTQADANTLQIAEIKEKIGTSAEDPV 1319  
QY 1321 SEVPVAVSHPRTKSRLOQSSLSAABEIGCNQTTQADANTLQIAEIKEKIGTSAEDPV 1380  
Db 1320 SEVPVAVSHPRTKSRLOQSSLSAABEIGCNQTTQADANTLQIAEIKEKIGTSAEDPV 1379  
QY 1381 PLMFRCSTVSASLDSFERSSTASSVQSEPCGMSVGIISPSDLDPSPGQTPPERSKTTP 1440  
Db 1380 PLMFRCSTVSASLDSFERSSTASSVQSEPCGMSVGIISPSDLDPSPGQTPPERSKTTP 1439  
QY 1441 PPPQTAQTKREVPKNAKPTAEKRESGPKQAAVNAVQVQVLPDADTLLHFAESTPDGF 1500  
Db 1440 PPPQTAQTKREVPKNAKPTAEKRESGPKQAAVNAVQVQVLPDADTLLHFAESTPDGF 1499  
QY 1501 SCSSSLSALSDDEPFIQKVELRIMPVQENDNGNETESEOPKESNENQKEAKTIDSE 1560  
Db 1500 SCSSSLSALSDDEPFIQKVELRIMPVQENDNGNETESEOPKESNENQKEAKTIDSE 1559  
QY 1561 KOLLDDSDDDDETEIEECIIISAMPTKSSRKAKKPAQATASKLPPFVARKPSQLPVYKLLPS 1620  
Db 1560 KOLLDDSDDDDETEIEECIIISAMPTKSSRKAKKPAQATASKLPPFVARKPSQLPVYKLLPS 1619  
QY 1621 QNRLOPQKHVSFTPGDDMPRVYCVGEGTPIINFSTATSLSDLTIESPPNELAAGEVGRGAQ 1680  
Db 1620 QNRLOPQKHVSFTPGDDMPRVYCVGEGTPIINFSTATSLSDLTIESPPNELAAGEVGRGAQ 1679  
QY 1681 SGEFEKRDITPEGRSTDEAQQGKTSVITIPELDDNKAEEGDILAEACINSAMPKGSHPK 1740  
Db 1680 SGEFEKRDITPEGRSTDEAQQGKTSVITIPELDDNKAEEGDILAEACINSAMPKGSHPK 1739  
QY 1741 FRVKIMDQVQASASSAPNKNQLDGKKKTPSPVKPIQNTYTRVRKNADSKNIN 1800  
Db 1740 FRVKIMDQVQASASSAPNKNQLDGKKKTPSPVKPIQNTYTRVRKNADSKNIN 1799  
QY 1801 AERFVSDNKSQKQNLKNSKDFNDKLPNNEDRVGSGFAFSPHYTPIEGTVCFSRND 1860  
Db 1800 AERFVSDNKSQKQNLKNSKDFNDKLPNNEDRVGSGFAFSPHYTPIEGTVCFSRND 1859  
QY 1861 SLSSLDFFDDDDVDLSREKAEALAKAKENKESAKVTSHTELTSNQOSANKTQAIKQPINR 1920  
Db 1860 SLSSLDFFDDDDVDLSREKAEALAKAKENKESAKVTSHTELTSNQOSANKTQAIKQPINR 1919  
QY 1921 GQKPILOKQSTFPOSSKDIIPRGAATDEKLNQFAIENTPVCFSNSSLSDIDQENN 1980  
Db 1920 GQKPILOKQSTFPOSSKDIIPRGAATDEKLNQFAIENTPVCFSNSSLSDIDQENN 1979  
QY 1981 NKENPPIKETEPDQSGEBSKPOAGYAPKSFHVEDTVCFSRNSLSLSIDSEDDLQ 2040  
Db 1980 NKENPPIKETEPDQSGEBSKPOAGYAPKSFHVEDTVCFSRNSLSLSIDSEDDLQ 2039  
QY 2041 ECISAMPKPKKPSRLKGDNEKHSRNMGGILGEDTLTDLKDIOQPDSEHGLSPDSENF 2100  
Db 2040 ECISAMPKPKKPSRLKGDNEKHSRNMGGILGEDTLTDLKDIOQPDSEHGLSPDSENF 2099  
QY 2101 WKAIQEGANSIYSSLHQAARAACTQOASSDSLSLKSGLSGSPHLPDQEEKPT 2160  
Db 2100 WKAIQEGANSIYSSLHQAARAACTQOASSDSLSLKSGLSGSPHLPDQEEKPT 2159  
QY 2161 SNKGPRILKPGKSTLTETKKIESESGIKGGKVKYKSLITGVRSNBSISGQMKQPLQAN 2220  
Db 2160 SNKGPRILKPGKSTLTETKKIESESGIKGGKVKYKSLITGVRSNBSISGQMKQPLQAN 2219  
QY 2221 MPISISGRWTWHPGVNRNSSSTSPVSKGPKLPKTPASKSPSEGOTATTSRCAKPSVKS 2280  
Db 2220 MPISISGRWTWHPGVNRNSSSTSPVSKGPKLPKTPASKSPSEGOTATTSRCAKPSVKS 2279  
QY 2281 ELSPVARQTSQIGSSKAPSRSGSRDSTPSRPAQPLSRPIQSPGNSISPGNGISPPN 2340  
Db 2280 ELSPVARQTSQIGSSKAPSRSGSRDSTPSRPAQPLSRPIQSPGNSISPGNGISPPN 2339

QY 2341 KLSQLPRTSPSTASTKSSGSGKSVYTSFGQMSQQNLTKQTGLSKNASSIPRSESASKG 2400  
Db 2340 KLSQLPRTSPSTASTKSSGSGKSVYTSFGQMSQQNLTKQTGLSKNASSIPRSESASKG 2399  
QY 2401 LNOMNNGNANKKVELSRMSSTKSSGSDRSEPRPVLVQSTFIKEAPSTLRKLEESA 2460  
Db 2400 LNOMNNGNANKKVELSRMSSTKSSGSDRSEPRPVLVQSTFIKEAPSTLRKLEESA 2459  
QY 2461 SFESLSPSSRPASPTSAQTPVLSPLDMSLSTHSSVQAGGWRKLPNLSPTIEYNDG 2520  
Db 2460 SFESLSPSSRPASPTSAQTPVLSPLDMSLSTHSSVQAGGWRKLPNLSPTIEYNDG 2519  
QY 2521 RPAKRHDIAARSHESPSRLPINRSGTWKREHSHSSSLPRVSTWRTGTSSSSILSASSES 2580  
Db 2520 RPAKRHDIAARSHESPSRLPINRSGTWKREHSHSSSLPRVSTWRTGTSSSSILSASSES 2579  
QY 2581 SERAKSEDEKHVNSISGTQSKENOVSAKGTWEKIKENBEPSTNSTSQTVSSGATNGAES 2640  
Db 2580 SERAKSEDEKHVNSISGTQSKENOVSAKGTWEKIKENBEPSTNSTSQTVSSGATNGAES 2639  
QY 2641 KTLIYQMAPAVSTEDVWVRIEDCPINNPRSGSPGTGNTPPVIDSVSEKANPNIKOSKDN 2700  
Db 2640 KTLIYQMAPAVSTEDVWVRIEDCPINNPRSGSPGTGNTPPVIDSVSEKANPNIKOSKDN 2699  
QY 2701 QAKQNVGNGSVPMRTVGLNRLNSFIQVDAPDQKTEIKPGQNNPVPVSETNESSIVERT 2760  
Db 2700 QAKQNVGNGSVPMRTVGLNRLNSFIQVDAPDQKTEIKPGQNNPVPVSETNESSIVERT 2759  
QY 2761 PFSSSSSSKHSSPSGTVAAARVTPFNYPNPSPRKSSADSTARPSPQIPTPVNNNTKCRDST 2820  
Db 2760 PFSSSSSSKHSSPSGTVAAARVTPFNYPNPSPRKSSADSTARPSPQIPTPVNNNTKCRDST 2819  
QY 2821 DSTESSGTQSPKHSYLSVTSV 2843  
Db 2820 DSTESSGTQSPKHSYLSVTSV 2842

RESULT 9  
ABG90968  
ID ABG90968 standard; peptide; 2842 AA.  
XX  
AC ABG90968;  
XX  
DT 29-NOV-2002 (first entry)  
XX Human APC protein.  
DE Adenomatous polyposis coli; APC; human; neoplastic tissue;  
KW mutation detection; tumour; cancer.  
XX Homo sapiens.  
OS  
XX US6413727-B1.  
XX  
PD 02-JUL-2002.  
XX  
PF 25-MAY-1995; 95US-00449731.  
XX  
PR 16-JAN-1991; 91GB-00000962.  
PR 16-JAN-1991; 91GB-00000963.  
PR 16-JAN-1991; 91GB-00000974.  
PR 16-JAN-1991; 91GB-00000975.  
PR 08-AUG-1991; 91US-00741940.  
PR 12-AUG-1994; 94US-00289548.  
XX (UKJO ) UNIV JOHNS HOPKINS.  
PA (UTAH ) UNIV UTAH.  
PA (NICA-) JAPANESE FOUND CANCER RES.  
PA (ZENE ) ZENECA LTD.  
XX Albertsen H, Anand R, Carlson M, Groden J, Hedge PJ, Joslyn G;  
PI Kinzler K, Markham AF, Nakamura Y, Thliveris A, Vogelstein B;  
PI White RL;



XX DR WPI; 2002-641559/69.

XX PT Method to aid in the diagnosis/prognosis of neoplastic tissues in humans,

XX PT by detecting somatic alteration of wild-type APC protein in tumor tissue

XX PT isolated from human, the alteration indicating neoplasia of the tissue.

XX PS Example 15; Fig 3; 140pp; English.

XX CC This invention relates to a novel method to aid in the diagnosis or

CC CC prognosis of a neoplastic tissue of a human. The method involves

CC CC detecting somatic alteration of wild-type adenomatous polyposis coli

CC CC protein in a tumour tissue isolated from a human (the alteration is

CC CC indicating neoplasia of the tissue). The method of the invention is

CC CC useful in diagnosis or prognosis of a neoplastic tissue of a human. the

CC CC method is useful in detection of genetic predisposition to cancer. The

CC CC present sequence represents a peptide sequence used in the method of the

XX CC invention

XX CC Sequence 2842 AA;

XX CC Query Match 99.9%; Score 14548.5; DB 5; Length 2842;

XX CC Best Local Similarity 100.0%; Pred. No. 0;

XX CC Matches 2842; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 MAAASYDQLLKQVEALKQVENSINLRQELSDNSNHLTKLETSANMKEVILKQLOQSIIDEAM 60

Db 1 MAAASYDQLLKQVEALKQVENSINLRQELSDNSNHLTKLETSANMKEVILKQLOQSIIDEAM 60

QY 61 ASSQOIDLLELRLKELNLDNSNFPQVKLRKXSLRSYSGREGSVSSSGECSVPVPMGSFPR 120

Db 61 ASSQOIDLLELRLKELNLDNSNFPQVKLRKXSLRSYSGREGSVSSSGECSVPVPMGSFPR 120

QY 121 RGFVNGRESTGYLELEKERSLLADLDKEEKEKDWYQAQLNLTNRIDSLPLTENFSL 180

Db 121 RGFVNGRESTGYLELEKERSLLADLDKEEKEKDWYQAQLNLTNRIDSLPLTENFSL 179

QY 181 QTDWTRQLEVEARQIRVAMEEQLGTCODMKEKAQRRIARIQIENKILRIIROLLOSOAT 240

Db 180 QTDWTRQLEVEARQIRVAMEEQLGTCODMKEKAQRRIARIQIENKILRIIROLLOSOAT 239

QY 241 EAERSSQNHETGSHDAERQEGGVGEINWATSGNGQGGSTTRMDHETASVLSSSSSTHSA 300

Db 240 EAERSSQNHETGSHDAERQEGGVGEINWATSGNGQGGSTTRMDHETASVLSSSSSTHSA 299

QY 301 PRLTSHLGTKEVWYSLLSMLGTHDMDSRLLTLLAMSSQSDSCISMRQSGCLPLLIQLL 360

Db 300 PRLTSHLGTKEVWYSLLSMLGTHDMDSRLLTLLAMSSQSDSCISMRQSGCLPLLIQLL 359

QY 361 HGNDKOSVLLGNRSGSKEARASAAALHNIHSGQDDKRRREIRVLLHLEQIRAYCETC 420

Db 360 HGNDKOSVLLGNRSGSKEARASAAALHNIHSGQDDKRRREIRVLLHLEQIRAYCETC 419

QY 421 WEWQEAHEPGMDQKNPMPAPVHQICPAVCVLMKLSFDBEHRHANNELGGLOAIAELLQ 480

Db 420 WEWQEAHEPGMDQKNPMPAPVHQICPAVCVLMKLSFDBEHRHANNELGGLOAIAELLQ 479

QY 481 VDCEMYGLTNDHYSITLRRVAGMALTNLTPGDVANKATILCSMKGCMREALVAOLKSESEDL 540

Db 480 VDCEMYGLTNDHYSITLRRVAGMALTNLTPGDVANKATILCSMKGCMREALVAOLKSESEDL 539

QY 541 QQVIAVLNRLSWRADVNSKTLREVGSVKALMECALEVKEKSTLKSVLKSLWNLNSAHT 600

Db 540 QQVIAVLNRLSWRADVNSKTLREVGSVKALMECALEVKEKSTLKSVLKSLWNLNSAHT 599

QY 601 ENKADICAVDGAFLVGLTYTSQNTLAIIBSGGGILRNYSLLIATNEDHROILLRENN 660

Db 600 ENKADICAVDGAFLVGLTYTSQNTLAIIBSGGGILRNYSLLIATNEDHROILLRENN 659

QY 661 CLOTLLQHLKSHSLTIVSNACGTLWNLNLSARNPKDQPALWDMGAVSMKXLIHSHKHQIAM 720

Db 660 CLOTLLQHLKSHSLTIVSNACGTLWNLNLSARNPKDQPALWDMGAVSMKXLIHSHKHQIAM 719

QY 721 GSAAALRNLMANRPAKYKDANIMSPGSLPSLHVRKQKALELDAOHLSETTFDNDNLS 780

Db 720 GSAAALRNLMANRPAKYKDANIMSPGSLPSLHVRKQKALELDAOHLSETTFDNDNLS 779

QY 781 PKASHRSQRHQRKQSLGYDVFDTNRHDDNRSDNFTNMTVLSPLYNTVLPSSSSSRGS 840

Db 780 PKASHRSQRHQRKQSLGYDVFDTNRHDDNRSDNFTNMTVLSPLYNTVLPSSSSSRGS 839

QY 841 LDSSSEKDRSLERERIGLGNYPATENPCTSKRGLOISTTAAQIAKVMVEYSATHTS 900

Db 840 LDSSSEKDRSLERERIGLGNYPATENPCTSKRGLOISTTAAQIAKVMVEYSATHTS 899

QY 901 QEDRSSGTTBELHCVTDERNALRSSAAHSTNFTYNTKSENSNRTCTMPYAKLEYKRSS 960

Db 900 QEDRSSGTTBELHCVTDERNALRSSAAHSTNFTYNTKSENSNRTCTMPYAKLEYKRSS 959

QY 961 NDSLNSVSSSGYGRGQMKPSIESYSEDDDESCKSYGOYPADLAHKTSHANHMDNDGE 1020

Db 960 NDSLNSVSSSGYGRGQMKPSIESYSEDDDESCKSYGOYPADLAHKTSHANHMDNDGE 1019

QY 1021 LDTPINYSLKYSDQLNSGRQSPQNERWARPKHIIEDIEIKQSEORQSRNQSTTYPVTE 1080

Db 1020 LDTPINYSLKYSDQLNSGRQSPQNERWARPKHIIEDIEIKQSEORQSRNQSTTYPVTE 1079

QY 1081 STDDKHLKFPHFQGCQCVSPYRSGANGSETNVRGNSHGINQNVSLQCEDDYEYDDKP 1140

Db 1080 STDDKHLKFPHFQGCQCVSPYRSGANGSETNVRGNSHGINQNVSLQCEDDYEYDDKP 1139

QY 1141 TNYSERVSEEEQHHEERPTNYSIKYNEEKRHVDQPIDYSLKYATDIPSSQKQSFSEKS 1200

Db 1140 TNYSERVSEEEQHHEERPTNYSIKYNEEKRHVDQPIDYSLKYATDIPSSQKQSFSEKS 1199

QY 1201 SSGQSKTEHMSSESSENTSTPSSNAKQNLHPSSAOSRSGQPOKAATCKVSSINQETIQ 1260

Db 1200 SSGQSKTEHMSSESSENTSTPSSNAKQNLHPSSAOSRSGQPOKAATCKVSSINQETIQ 1259

QY 1261 TYCVEDTPICFRCSLSLSLSAEDIEGCTTQDEASANTLOIABEIKIGTSAEDPV 1320

Db 1260 TYCVEDTPICFRCSLSLSLSAEDIEGCTTQDEASANTLOIABEIKIGTSAEDPV 1319

QY 1321 SEVPAVSGHPPTKSRLQGSLSSESARHKAVERPSSGAKSPSKGAQTPKSPPEHYVQET 1380

Db 1320 SEVPAVSGHPPTKSRLQGSLSSESARHKAVERPSSGAKSPSKGAQTPKSPPEHYVQET 1379

QY 1381 PLMFSRCTSVSSLSDFSRSTASSVQSEPCSGMGVGIISPSDLPDPSGQTMPPSRKTPP 1440

Db 1380 PLMFSRCTSVSSLSDFSRSTASSVQSEPCSGMGVGIISPSDLPDPSGQTMPPSRKTPP 1439

QY 1441 PPPOTAQTKREVPKXKAPTAEKRESGPKQAAVNAVQVQLPDADTLLHFATESTPDGF 1500

Db 1440 PPPOTAQTKREVPKXKAPTAEKRESGPKQAAVNAVQVQLPDADTLLHFATESTPDGF 1499

QY 1501 SCSSLSLSALDDEFFIOKDVLELRIMPPVOENDNGNETESEQPKESNENQBEAKTIDSE 1560

Db 1500 SCSSLSLSALDDEFFIOKDVLELRIMPPVOENDNGNETESEQPKESNENQBEAKTIDSE 1559

QY 1561 KDLLDDSDDDDDIEILEECIIISAMPTKSRKAKKPAQATSKLPPPVARKPQLPVYKLLPS 1620

Db 1560 KDLLDDSDDDDDIEILEECIIISAMPTKSRKAKKPAQATSKLPPPVARKPQLPVYKLLPS 1619

QY 1621 QNRLOPKHVSFTPGDDMPRVYCVGEPTPIPFSTATSLDITSPSPNELAAGEVGRGAQ 1680

Db 1620 QNRLOPKHVSFTPGDDMPRVYCVGEPTPIPFSTATSLDITSPSPNELAAGEVGRGAQ 1679

QY 1681 SGEFEKDTTPTGSRSTDEAGGKTSSVTIPELDDNKAEBGDIILAEICINSAMPKGSHPK 1740

Db 1680 SGEFEKDTTPTGSRSTDEAGGKTSSVTIPELDDNKAEBGDIILAEICINSAMPKGSHPK 1739

QY 1741 FRVKIKIMQVQQAASASAPNKNQLDGKKKKPTSPVKPIPONTTEYTRVRKNADSKNNLN 1800

Db 1740 FRVKIKIMQVQQAASASAPNKNQLDGKKKKPTSPVKPIPONTTEYTRVRKNADSKNNLN 1799

QY 1801 AERVFSNDKSKQNLKNNSKDFNDKLPNNEDRVGRSFAFDSPHHYTPIEGTPTCYCFERN 1860



DB 1800 AERFSDNKKQNLKNNKDFNDKLFNNEDVRGSAFDPHYTPIEGTPTCFGRND 1859  
 QY 1861 SLSSLDFFDDVLDLREKAEALREKAKENKESAKVTSHTELTNSQOSANKTOAIKQPINR 1920  
 DB 1860 SLSSLDFFDDVLDLREKAEALREKAKENKESAKVTSHTELTNSQOSANKTOAIKQPINR 1919  
 QY 1921 GQPKPILQKQSTFPOSSXDIIDRGAATDEKLNFAIENTPVCFHNSLSLSIDQENN 1980  
 DB 1920 GQPKPILQKQSTFPOSSXDIIDRGAATDEKLNFAIENTPVCFHNSLSLSIDQENN 1979  
 QY 1981 NKENEPIKETPPDSCGPPSPKQASGYAPKPFHVEDTPVCFSRNSSLSSLSIDSEDDLLQ 2040  
 DB 1980 NKENEPIKETPPDSCGPPSPKQASGYAPKPFHVEDTPVCFSRNSSLSSLSIDSEDDLLQ 2039  
 QY 2041 ECISAMPKPKKPSRLKGDNEKHSPPNMGILGEDLTLDLKDIOQPDSEHGLSPDSENF 2100  
 DB 2040 ECISAMPKPKKPSRLKGDNEKHSPPNMGILGEDLTLDLKDIOQPDSEHGLSPDSENF 2099  
 QY 2101 WKAIQEGANSIVSSLHQAAAACLSQASSDSDSLSLKSGISLSPHLPDQEEKPF 2160  
 DB 2100 WKAIQEGANSIVSSLHQAAAACLSQASSDSDSLSLKSGISLSPHLPDQEEKPF 2159  
 QY 2161 SNKGRPILKPKGKSTLETKKIESEKGIKGGKVKYKSLITGKVRNSNISISQMKQPLQAN 2220  
 DB 2160 SNKGRPILKPKGKSTLETKKIESEKGIKGGKVKYKSLITGKVRNSNISISQMKQPLQAN 2219  
 QY 2221 MFSISGRGTMHIIPGVNRSSSTSPVSKGPPPLKTPASKSPSEGGTATTSRGAKPVS 2280  
 DB 2220 MFSISGRGTMHIIPGVNRSSSTSPVSKGPPPLKTPASKSPSEGGTATTSRGAKPVS 2279  
 QY 2281 ELSPVARTQSLGGSKKAPSRSGSDSTPSRPAQOPLSRPTQSPORNISIPGRNGISPPN 2340  
 DB 2280 ELSPVARTQSLGGSKKAPSRSGSDSTPSRPAQOPLSRPTQSPORNISIPGRNGISPPN 2339  
 QY 2341 KLSQLPRTSSPTASTKSGSGKMYTSPGRMSQOQLTKOTGLSKNASSIPRESASKG 2400  
 DB 2340 KLSQLPRTSSPTASTKSGSGKMYTSPGRMSQOQLTKOTGLSKNASSIPRESASKG 2399  
 QY 2401 LQNMANGANKKVLRSMSSTKSGSESDSERPVLVQSTFIKEAPSPILRRKLEESA 2460  
 DB 2400 LQNMANGANKKVLRSMSSTKSGSESDSERPVLVQSTFIKEAPSPILRRKLEESA 2459  
 QY 2461 SPESLSPSRPASPTRSQATPVLSPSLPDMSLSTHSSVQAGGWRKLPNLSPTIEYNDG 2520  
 DB 2460 SPESLSPSRPASPTRSQATPVLSPSLPDMSLSTHSSVQAGGWRKLPNLSPTIEYNDG 2519  
 QY 2521 RPAKHDIARSHSPSRPLPNRSGTWKREHSHKSSSLPRVSTWRTGSSSILSASSES 2580  
 DB 2520 RPAKHDIARSHSPSRPLPNRSGTWKREHSHKSSSLPRVSTWRTGSSSILSASSES 2579  
 QY 2581 SEKAKSEDEKHVNSISGTQSKXENOVSAKGTWRKIKENEFSPNTSOTVSSGATNGAES 2640  
 DB 2580 SEKAKSEDEKHVNSISGTQSKXENOVSAKGTWRKIKENEFSPNTSOTVSSGATNGAES 2639  
 QY 2641 KTLIYQMAPAVSKTEDVWVRIEDCPINNPRSGRSPGTNTPPVIDSVSEKANPNIKDSKDN 2700  
 DB 2640 KTLIYQMAPAVSKTEDVWVRIEDCPINNPRSGRSPGTNTPPVIDSVSEKANPNIKDSKDN 2699  
 QY 2701 QAKQNVGSGVPMRTVGLNRLNSFTOVDAPQKGTETKPGONNPVPVSETHNESSIVERT 2760  
 DB 2700 QAKQNVGSGVPMRTVGLNRLNSFTOVDAPQKGTETKPGONNPVPVSETHNESSIVERT 2759  
 QY 2761 PFSSSSSSKHSPSGTVAARVTPFNYPNPSFRKSSADSTARSQIPTPVNNNTKKRDSKT 2820  
 DB 2760 PFSSSSSSKHSPSGTVAARVTPFNYPNPSFRKSSADSTARSQIPTPVNNNTKKRDSKT 2819  
 QY 2821 DSTESSGTSQPKRHSGSYLTVSV 2843  
 DB 2820 DSTESSGTSQPKRHSGSYLTVSV 2842

RESULT 10

AAW11922  
 ID AAW11922 standard; protein; 2843 AA.  
 AC AAW11922;  
 DT 07-MAY-1997 (first entry)  
 DE Adenomatous polyposis coli protein.  
 XX Adenomatous polyposis coli; APC; human; antibody; mutant detection.  
 OS Homo sapiens.  
 PN JP06347459-A.  
 PD 22-DEC-1994.  
 PF 07-JUN-1993; 93JP-00136102.  
 PR 07-JUN-1993; 93JP-00136102.  
 FA (SANK-) SANKO JUNYAKU CO LTD.  
 XX WPI; 1995-070957/10.  
 PT Detection of abnormal human APC (adenomatous polyposis coli) gene - using antibodies directed against the APC gene product.  
 PS Claim 1; Page 7-13; 13pp; Japanese.  
 CC This sequence represents the wild type human adenomatous polyposis coli (APC) protein. This sequence, and the APC peptide fragments represented by AAW11923-W11929 can be used in the method of the invention. The method of the invention is for the detection of abnormal APC, using antibodies directed against this sequence, or one of the partial peptide sequences  
 SQ Sequence 2843 AA;  
 Query Match 99.9%; Score 14546; DB 2; Length 2843;  
 Best Local Similarity 99.9%; Pred. No. 0;  
 Matches 2840; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
 QY 1 MAAASYDQLLKQVEALKWENSNLRQELDENSNLTKLETSANNKEVVKLOQSIIEDEAM 60  
 DB 1 MAAASYDQLLKQVEALKWENSNLRQELDENSNLTKLETSANNKEVVKLOQSIIEDEAM 60  
 QY 61 ASSQIDLLERLKEINLDSSNFFGVKLRKMSLRSGVSGREGSVSSRSGECPVPMGSFPR 120  
 DB 61 ASSQIDLLERLKEINLDSSNFFGVKLRKMSLRSGVSGREGSVSSRSGECPVPMGSFPR 120  
 QY 121 RGFVNGSRESTGYLEELKERSLLADLDKEEKEKWYQAQLNLTNRIDSLPTENFSL 180  
 DB 121 RGFVNGSRESTGYLEELKERSLLADLDKEEKEKWYQAQLNLTNRIDSLPTENFSL 180  
 QY 181 QTDMTRQLEYEARQIRVAMEEQGTCCDMKEKRAQRIARIQIQIEKDILRIQLLOQAT 240  
 DB 181 QTDMTRQLEYEARQIRVAMEEQGTCCDMKEKRAQRIARIQIQIEKDILRIQLLOQAT 240  
 QY 241 EAERSSQNKHETGSHDAERQNEGQVGEINWATSGNQGSTTRMDHETASVSSSTHSA 300  
 DB 241 EAERSSQNKHETGSHDAERQNEGQVGEINWATSGNQGSTTRMDHETASVSSSTHSA 300  
 QY 301 PRRLTSHLGTKEVMVYSLLSMLGTHDKDDMSRILLAMSSQSDSCISNRQSGCPLLIQLL 360  
 DB 301 PRRLTSHLGTKEVMVYSLLSMLGTHDKDDMSRILLAMSSQSDSCISNRQSGCPLLIQLL 360  
 QY 361 HGNDKDSVLLGNRSGSEKARASAAALHNIHSOPDDKRGRRREIRVHLLEQIRAYCETC 420  
 DB 361 HGNDKDSVLLGNRSGSEKARASAAALHNIHSOPDDKRGRRREIRVHLLEQIRAYCETC 420  
 QY 421 WEQEAHEPGMDQKNPMPAPVEHQICPAVCVLMKLSFDEEHRHANNELGGLQIAELLQ 480  
 DB 421 WEQEAHEPGMDQKNPMPAPVEHQICPAVCVLMKLSFDEEHRHANNELGGLQIAELLQ 480

QY 481 VDCWYGLTNDHYGILTRRYAGMALTNLTFGDVANKATLCMSKGCMPALVAQLKSSSEDL 540  
 DB 481 VDCWYGLTNDHYGILTRRYAGMALTNLTFGDVANKATLCMSKGCMPALVAQLKSSSEDL 540  
 QY 541 QOVVASVLRLNLSWADVNSKTLREVGSVKALMECALEVKESTLKSVALNLSAHT 600  
 DB 541 QOVVASVLRLNLSWADVNSKTLREVGSVKALMECALEVKESTLKSVALNLSAHT 600  
 QY 601 ENKADI CAVDGALAFVLGTLTYRSQNTLAIIESGGGILRVNSSLIATNEDHQILRENN 660  
 DB 601 ENKADI CAVDGALAFVLGTLTYRSQNTLAIIESGGGILRVNSSLIATNEDHQILRENN 660  
 QY 661 CLQTLQHLKSHSLTIVSNACGTLWNLNARSNPQDEALWDMGAVSMLKNLIHSHKWMIAM 720  
 DB 661 CLQTLQHLKSHSLTIVSNACGTLWNLNARSNPQDEALWDMGAVSMLKNLIHSHKWMIAM 720  
 QY 721 GSAAALRLNMANRPAKYKDANIMSPGSLPSLHVKKQKALEAEIDAQHLSETFDNDNLNLS 780  
 DB 721 GSAAALRLNMANRPAKYKDANIMSPGSLPSLHVKKQKALEAEIDAQHLSETFDNDNLNLS 780  
 QY 781 PKASHRSKQRLKOSLYGQVDFDTRNDRDNDNPNNTGNMTVLPSPYLNNTVLPSSSSSRGS 840  
 DB 781 PKASHRSKQRLKOSLYGQVDFDTRNDRDNDNPNNTGNMTVLPSPYLNNTVLPSSSSSRGS 840  
 QY 841 LDRSRSEKRLERERGI GLGNHYHPATENPGTSSKRGLOISTTAAQIAKYMEEVSAIHTS 900  
 DB 841 LDRSRSEKRLERERGI GLGNHYHPATENPGTSSKRGLOISTTAAQIAKYMEEVSAIHTS 900  
 QY 901 QEDRSSGTTTELHCVTDERNALRSSAAHSTHNTYNTFKSENSNRCTCSMPYAKLEYKRSS 960  
 DB 901 QEDRSSGTTTELHCVTDERNALRSSAAHSTHNTYNTFKSENSNRCTCSMPYAKLEYKRSS 960  
 QY 961 NDSLSNVSSDGYKRGOMKPSIESYEDDESFCYGOYPADLAHKIHSANHMDNDGE 1020  
 DB 961 NDSLSNVSSDGYKRGOMKPSIESYEDDESFCYGOYPADLAHKIHSANHMDNDGE 1020  
 QY 1021 LDTPIYNSLYKSDQLNSGRQSPQNERWARPKHIIIEDEIKQSEQRNQSTTYPVYTE 1080  
 DB 1021 LDTPIYNSLYKSDQLNSGRQSPQNERWARPKHIIIEDEIKQSEQRNQSTTYPVYTE 1080  
 QY 1081 STDDKHLKFOHFCQBCVCPYSRGANGSETNRVGSNHNINONVSQSLCOEDDYEDDKP 1140  
 DB 1081 STDDKHLKFOHFCQBCVCPYSRGANGSETNRVGSNHNINONVSQSLCOEDDYEDDKP 1140  
 QY 1141 TNSYERYSEBEQHEERPTNYSIKYNEEKRVHDQPIDYSLKYATDIPSSQKQSFSSKS 1200  
 DB 1141 TNSYERYSEBEQHEERPTNYSIKYNEEKRVHDQPIDYSLKYATDIPSSQKQSFSSKS 1200  
 QY 1201 SSGOSSKTEHWSSENSTSPSSNAKQNLHPSSAQSRGQPKAATCKVSSINOETIQ 1260  
 DB 1201 SSGOSSKTEHWSSENSTSPSSNAKQNLHPSSAQSRGQPKAATCKVSSINOETIQ 1260  
 QY 1261 TYCVEDTPICFSRCSLSLSLSAEDEIGCNQTTQEADSANTLQIAIEIKIGITRSAEDPV 1320  
 DB 1261 TYCVEDTPICFSRCSLSLSLSAEDEIGCNQTTQEADSANTLQIAIEIKIGITRSAEDPV 1320  
 QY 1321 SEVPAVQHPRTKSSRLQGSLSLSSESARHKAFFSSGAKSPSKSGAQTPKSPPEHYVQET 1380  
 DB 1321 SEVPAVQHPRTKSSRLQGSLSLSSESARHKAFFSSGAKSPSKSGAQTPKSPPEHYVQET 1380  
 QY 1381 PLMFSRCTSVSSLDSPFSRSLIASSVQSEPCSGMVSGIIPSDLPDPGOTMPPSRSKTPP 1440  
 DB 1381 PLMFSRCTSVSSLDSPFSRSLIASSVQSEPCSGMVSGIIPSDLPDPGOTMPPSRSKTPP 1440  
 QY 1441 PPFQTAQTKREVPKNKAPTAEKESGPKQAANVAAYORVQVLPDADTLHLHFAESTPDGF 1500  
 DB 1441 PPFQTAQTKREVPKNKAPTAEKESGPKQAANVAAYORVQVLPDADTLHLHFAESTPDGF 1500  
 QY 1501 SCSSLSALSLEDPFFIQKQVELRIMPVQNDNGNETSEOPKESNENOEKAEKTI DSE 1560  
 DB 1501 SCSSLSALSLEDPFFIQKQVELRIMPVQNDNGNETSEOPKESNENOEKAEKTI DSE 1560

QY 1561 KOLLDDSDDDDIIEIEBECIIISAMPTKSSRKAKKPAQATASKLPPPVARKPSOLPVYKLLPS 1620  
 DB 1561 KOLLDDSDDDDIIEIEBECIIISAMPTKSSRKAKKPAQATASKLPPPVARKPSOLPVYKLLPS 1620  
 QY 1621 QNRLQFQKHVSFTPDGDDMPRVYCVVEGTPIINFSTATSLDLTIESPPNLAAGEVGRGAQ 1680  
 DB 1621 QNRLQFQKHVSFTPDGDDMPRVYCVVEGTPIINFSTATSLDLTIESPPNLAAGEVGRGAQ 1680  
 QY 1681 SGEFEKRTDITPEGRSTDEAOGGKTSSVTIPELDDNKAEBGDIILAEICINSAMPKGSHP 1740  
 DB 1681 SGEFEKRTDITPEGRSTDEAOGGKTSSVTIPELDDNKAEBGDIILAEICINSAMPKGSHP 1740  
 QY 1741 FRVKIMQVQOASASSAPNKNQLDGKKKPTSPVKPIPONTEYRTRVRKNADSKNLN 1800  
 DB 1741 FRVKIMQVQOASASSAPNKNQLDGKKKPTSPVKPIPONTEYRTRVRKNADSKNLN 1800  
 QY 1801 AERVFSDNKKQNLKXNSKDFNDKLPNNEDRVRGSAFDSPHHYTPIEGTPTCYFSRND 1860  
 DB 1801 AERVFSDNKKQNLKXNSKDFNDKLPNNEDRVRGSAFDSPHHYTPIEGTPTCYFSRND 1860  
 QY 1861 SLSSLDFFDDDDVLSREKAEIRKAKENKESAKVTSHTELTSNOOSANKTQAIKQPINR 1920  
 DB 1861 SLSSLDFFDDDDVLSREKAEIRKAKENKESAKVTSHTELTSNOOSANKTQAIKQPINR 1920  
 QY 1921 GQPKPILOKQSTFPQSSKDIIPDRGAATDEKLQNFALIENTPVCFSHNSLSLSLSDIDQENN 1980  
 DB 1921 GQPKPILOKQSTFPQSSKDIIPDRGAATDEKLQNFALIENTPVCFSHNSLSLSLSDIDQENN 1980  
 QY 1981 NKENEPIKETETPPDSQGEPSKQASGVAPKSFHVEDTPVCFSRNSLSLSLSDIDQENLLQ 2040  
 DB 1981 NKENEPIKETETPPDSQGEPSKQASGVAPKSFHVEDTPVCFSRNSLSLSLSDIDQENLLQ 2040  
 QY 2041 ECISAMPKKKPKRSLKGDNEKHSPRNMGIGILGEDTLDLKDIOQRPDSEHGLSPDSNF 2100  
 DB 2041 ECISAMPKKKPKRSLKGDNEKHSPRNMGIGILGEDTLDLKDIOQRPDSEHGLSPDSNF 2100  
 QY 2101 WKATOEGANSIVSSLHQAAAAACLSROASSDSISLXSGISILGSPFHLTPDQEEKPFT 2160  
 DB 2101 WKATOEGANSIVSSLHQAAAAACLSROASSDSISLXSGISILGSPFHLTPDQEEKPFT 2160  
 QY 2161 SNKGPRILLKPEKSTLETTKIESKGIKGGKVKYKSLITGKVESNSGEISQMKQKQLOAN 2220  
 DB 2161 SNKGPRILLKPEKSTLETTKIESKGIKGGKVKYKSLITGKVESNSGEISQMKQKQLOAN 2220  
 QY 2221 MFSISRGRTMIHIFGVNRNSSSTSPVSKGPKLTPASKSPSEGTATTPRGAKPSVKS 2280  
 DB 2221 MFSISRGRTMIHIFGVNRNSSSTSPVSKGPKLTPASKSPSEGTATTPRGAKPSVKS 2280  
 QY 2281 ELSVPARTSOIGSSKAPSPSGSRSDSTPSRPAOQPLSRPIQSPGRNSISPGRNGISPPN 2340  
 DB 2281 ELSVPARTSOIGSSKAPSPSGSRSDSTPSRPAOQPLSRPIQSPGRNSISPGRNGISPPN 2340  
 QY 2341 KLSQLPRTSSPTASTKSSSGSKMSYTSFGQMSQQNLTKQTGLSKNASSIPRSESASKG 2400  
 DB 2341 KLSQLPRTSSPTASTKSSSGSKMSYTSFGQMSQQNLTKQTGLSKNASSIPRSESASKG 2400  
 QY 2401 LNQMNGNANGANKVELSRMSTKSSGSEDSERBPVLVRSQTFIKEAPSTFLRKLBEESA 2460  
 DB 2401 LNQMNGNANGANKVELSRMSTKSSGSEDSERBPVLVRSQTFIKEAPSTFLRKLBEESA 2460  
 QY 2461 SPESLSPSSRPASTRSQAQTPVLSPLPMSLSTHSSVQAGGWRKLPNLSPTIEYNDG 2520  
 DB 2461 SPESLSPSSRPASTRSQAQTPVLSPLPMSLSTHSSVQAGGWRKLPNLSPTIEYNDG 2520  
 QY 2521 RPAKRHDIAHSHSESPRLPINRSGTWKRBHSHKSSSLPRVSTWRTGTGSSSSILSASSES 2580  
 DB 2521 RPAKRHDIAHSHSESPRLPINRSGTWKRBHSHKSSSLPRVSTWRTGTGSSSSILSASSES 2580  
 QY 2581 SEKAKSEDEKHNISIGTKOSKENOVSAKGTWRKIKENEFSPNTNSTQTVSSGATNCAES 2640  
 DB 2581 SEKAKSEDEKHNISIGTKOSKENOVSAKGTWRKIKENEFSPNTNSTQTVSSGATNCAES 2640  
 QY 2641 KTLIYQWAPAVSKTDEWVRIEDCPINNPRSGRSPTGNTPPVIDSVSEKANPNIKOSKDN 2700

Db 2641 KTLIYQMAVSKTEDVWRIEDCINNPRSGRTGNTFPVIDSVSEKANPNIKDSKN 2700  
QY 2701 QAKQVNGSVPMRTVGLNRLNSIQVDAPQKTEIKPGQNNPVPVSENESSIVERT 2760  
Db 2701 QAKQVNGSVPMRTVGLNRLNSIQVDAPQKTEIKPGQNNPVPVSENESSIVERT 2760  
QY 2761 PFSSSSSKHSGSPGTVAAARVTPFNPNPRKSSADSTSARPSQIPTPVNNTKKRDSKT 2820  
Db 2761 PFSSSSSKHSGSPGTVAAARVTPFNPNPRKSSADSTSARPSQIPTPVNNTKKRDSKT 2820  
QY 2821 DTESSSGTQSPKRHSGSYLVTSV 2843  
Db 2821 DTESSSGTQSPKRHSGSYLVTSV 2843

RESULT 11

AD65846  
ID AD65846 standard; protein; 2843 AA.

XX AC AD65846;

DT 29-JAN-2004 (first entry)

XX Human adenomatosis polyposis coli protein.

DE Human; adenomatosis polyposis coli protein; protein-protein interaction;  
KW protein array; PDZ domain; drug target screening.

XX OS Homo sapiens.

XX US2003170723-A1.

PN 11-SEP-2003.

XX 06-MAR-2002; 2002US-00092138.

PF 06-MAR-2002; 2002US-00092138.

XX (SATO/) SATO T.

PA Sato T;

XX WPI; 2003-852032/79.

XX Preparing a protein array useful for screening drug targets comprises  
PT depositing an array of a first protein on substrate, and applying a  
PT second protein comprising an amino acid sequence that binds to a domain  
PT of the first protein.

XX Disclosure; SEQ ID NO 30; 60pp; English.

PS The invention relates to a method for preparing a protein array based on  
CC protein-protein interaction, by depositing an array of a first protein  
CC comprising a PDZ domain on a substrate, and applying a second protein  
CC comprising an amino acid sequence that binds to the PDZ domain of the  
CC first protein. The method is useful for preparing protein arrays based on  
CC biochemical protein-protein interactions. Arrays produced by this method  
CC are useful for screening drug targets. This sequence represents the human  
CC adenomatosis polyposis coli protein, used in the method of the invention.

XX Sequence 2843 AA;

QY Query Match 99.8%; Score 14539; DB 7; Length 2843;

Best Local Similarity 99.8%; Pred. No. 0;  
Matches 2837; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 MAAASYDQLLKQVEALKMENSRLRQLEEDNSNHLTKLETEASNMKVEVLKQLQGSIEDAM 60

Db 1 MAAASYDQLLKQVEALKMENSRLRQLEEDNSNHLTKLETEASNMKVEVLKQLQGSIEDAM 60

QY 61 ASSGQIDLLERLKEALNLDSSNPFQVKLRSKWLSRYSGRSGSVSSRSGECSPVPMGSPFR 120

Db 61 ASSGQIDLLERLKEALNLDSSNPFQVKLRSKWLSRYSGRSGSVSSRSGECSPVPMGSPFR 120  
QY 121 RGFVNGSRSTGYLELEKERSILLADLDKEEKEKWYQAQLQNLTKRIDSLPLTENFSL 180  
Db 121 RGFVNGSRSTGYLELEKERSILLADLDKEEKEKWYQAQLQNLTKRIDSLPLTENFSL 180  
QY 181 QTDMTROQLYEARQIRVAMEEOLGTCDMEKAKQRIARIQOIEKIDILIRIQLLOSOAT 240  
Db 181 QTDMTROQLYEARQIRVAMEEOLGTCDMEKAKQRIARIQOIEKIDILIRIQLLOSOAT 240  
QY 241 EAEASSQNKHETGSHDAERONEGGVGEINMATSNGQGGSTTRMDHETASVLSSTSTHSA 300  
Db 241 EAEASSQNKHETGSHDAERONEGGVGEINMATSNGQGGSTTRMDHETASVLSSTSTHSA 300  
QY 301 PRRLTSHLGTKEWVYSLLSVLGTHDKDDMSRTLLAMSSQDSCISMRSQGCPLLIQLL 360  
Db 301 PRRLTSHLGTKEWVYSLLSVLGTHDKDDMSRTLLAMSSQDSCISMRSQGCPLLIQLL 360  
QY 361 HGNDKDSVLLNGSRGSKAARASAAALHNIHSPQDDKRGREIRVHLHLEQIRAYCETC 420  
Db 361 HGNDKDSVLLNGSRGSKAARASAAALHNIHSPQDDKRGREIRVHLHLEQIRAYCETC 420  
QY 421 WEQEAEHPGMDQKNPMPAPVEHOICPAVCVLMKLSFDEEHRHAMNELGGLQIAELLQ 480  
Db 421 WEQEAEHPGMDQKNPMPAPVEHOICPAVCVLMKLSFDEEHRHAMNELGGLQIAELLQ 480  
QY 481 VDCEMYGLTNDHYSITLRRYAGMALTNLTFDGVANKATLCSMKGCMREALVAQLKSESED 540  
Db 481 VDCEMYGLTNDHYSITLRRYAGMALTNLTFDGVANKATLCSMKGCMREALVAQLKSESED 540  
QY 541 QQVIASVLRNLSWRADVNSKTLREVGSVKALMECALEVKEKSTLKSVLKALNLSAHT 600  
Db 541 QQVIASVLRNLSWRADVNSKTLREVGSVKALMECALEVKEKSTLKSVLKALNLSAHT 600  
QY 601 ENKADICAVDGAFLVGTLYTYSQNTLAIISGGGILRNVSLLIATNEDHRIQLRENN 660  
Db 601 ENKADICAVDGAFLVGTLYTYSQNTLAIISGGGILRNVSLLIATNEDHRIQLRENN 660  
QY 661 CLQTLLOHLKSHSLTIVSNACGLTNLSARNPKDQALMDGAVSMKLIHSHKHVIAM 720  
Db 661 CLQTLLOHLKSHSLTIVSNACGLTNLSARNPKDQALMDGAVSMKLIHSHKHVIAM 720  
QY 721 GSAAALENMANRPKDYKDNIMSPGSSLPKLVKOKALELDAHLSETFDNDLS 780  
Db 721 GSAAALENMANRPKDYKDNIMSPGSSLPKLVKOKALELDAHLSETFDNDLS 780  
QY 781 PKASHRSKQHKQSLYGDYVFDNRHDDNRSDFNTGNMTVLPYLNTTVLPSSSSSRGS 840  
Db 781 PKASHRSKQHKQSLYGDYVFDNRHDDNRSDFNTGNMTVLPYLNTTVLPSSSSSRGS 840  
QY 841 LDSRSRSEKDRSLERERGIGLGNYPATENPGTSKRGLOISTTAAQIAKWEVSAIHTS 900  
Db 841 LDSRSRSEKDRSLERERGIGLGNYPATENPGTSKRGLOISTTAAQIAKWEVSAIHTS 900  
QY 901 QEDRSSSTTELHCVTDERNALRRSSAAHTSNFTKSENRTCSMPYAKLEYKRSS 960  
Db 901 QEDRSSSTTELHCVTDERNALRRSSAAHTSNFTKSENRTCSMPYAKLEYKRSS 960  
QY 961 NDSLNSVSSSDGYGKQGMKPSIESYSEDESFCGYQPADLAHKIHSANHMDNDGE 1020  
Db 961 NDSLNSVSSSDGYGKQGMKPSIESYSEDESFCGYQPADLAHKIHSANHMDNDGE 1020  
QY 1021 LDTPINYSKYSDQNLNSGQSPQNERWAPKHIIEDEIKQSEQRQSRNQSTTYPVTE 1080  
Db 1021 LDTPINYSKYSDQNLNSGQSPQNERWAPKHIIEDEIKQSEQRQSRNQSTTYPVTE 1080  
QY 1081 STDDKHLKFPHFQGCQECVSPYRSRGANGSETNRVGSNHGINQVQSQCQEDDYEDDKP 1140  
Db 1081 STDDKHLKFPHFQGCQECVSPYRSRGANGSETNRVGSNHGINQVQSQCQEDDYEDDKP 1140  
QY 1141 TNSYSEYSEEOHEEERPTNYSIKYNEEKHVDPQIDYSLKYATDIPSSKQSFSEKS 1200  
Db 1141 TNSYSEYSEEOHEEERPTNYSIKYNEEKHVDPQIDYSLKYATDIPSSKQSFSEKS 1200

QY 1201 SSGOSSKTEHMSSESSENTSTPSSNAKRONLHPSSAQSRSGQPOKAATCKVSSINQETIQ 1260  
 Db 1201 SSGOSSKTEHMSSESSENTSTPSSNAKRONLHPSSAQSRSGQPOKAATCKVSSINQETIQ 1260  
 QY 1261 TYCVEDTPICFRSCSSUSSLSAEDIEICNQTQOADSANTLOIABIKKIGITRSEADPV 1320  
 Db 1261 TYCVEDTPICFRSCSSUSSLSAEDIEICNQTQOADSANTLOIABIKKIGITRSEADPV 1320  
 QY 1321 SEVPAVQCHPRTKSSRLQGGSSLSAEDIEICNQTQOADSANTLOIABIKKIGITRSEADPV 1380  
 Db 1321 SEVPAVQCHPRTKSSRLQGGSSLSAEDIEICNQTQOADSANTLOIABIKKIGITRSEADPV 1380  
 QY 1381 PLMFSRCTSVSSLDSPFSRGIASVQSEPCGMVSGIISPSDLPDPSGQTMPPSRSKTTP 1440  
 Db 1381 PLMFSRCTSVSSLDSPFSRGIASVQSEPCGMVSGIISPSDLPDPSGQTMPPSRSKTTP 1440  
 QY 1441 PPPQATQKREVPKNAKAPTAEKRESGPKQAANAQVORVQVLPDADTLHFAESTPDGF 1500  
 Db 1441 PPPQATQKREVPKNAKAPTAEKRESGPKQAANAQVORVQVLPDADTLHFAESTPDGF 1500  
 QY 1501 SCSSLSALSIDEFFIQKVELRIMPVQENDNGNETESQPKESNENQEKAEKIDSE 1560  
 Db 1501 SCSSLSALSIDEFFIQKVELRIMPVQENDNGNETESQPKESNENQEKAEKIDSE 1560  
 QY 1561 KOLLDDSDDDIEILEBECIISAMPTKSSRAKPAQATASKLPPVAKPSQLPVYKLLPS 1620  
 Db 1561 KOLLDDSDDDIEILEBECIISAMPTKSSRAKPAQATASKLPPVAKPSQLPVYKLLPS 1620  
 QY 1621 QNRLOPKHVSFTGDDMPRVYCEVGTPIINFSTATSLSDLTIESPNELAAGSVRGGAQ 1680  
 Db 1621 QNRLOPKHVSFTGDDMPRVYCEVGTPIINFSTATSLSDLTIESPNELAAGSVRGGAQ 1680  
 QY 1681 SGEFEKRDITTEGRSTDEAOGGKTSVTTPELDDNKAEGDILABEINSAMPKGSHP 1740  
 Db 1681 SGEFEKRDITTEGRSTDEAOGGKTSVTTPELDDNKAEGDILABEINSAMPKGSHP 1740  
 QY 1741 FRVKKIMDQVQOASASSANKNOLDGKKKPTSPVKPIQNTYEYRTRVRKNADSKNNLN 1800  
 Db 1741 FRVKKIMDQVQOASASSANKNOLDGKKKPTSPVKPIQNTYEYRTRVRKNADSKNNLN 1800  
 QY 1801 AERFSDNKKQKQNLKNSKDFNDKLPNNEDVRGSGFAFDSPHHYTPIEGTFCFSRND 1860  
 Db 1801 AERFSDNKKQKQNLKNSKDFNDKLPNNEDVRGSGFAFDSPHHYTPIEGTFCFSRND 1860  
 QY 1861 SLSSLDFFDDDDVLSREKAEIRAKENKESFAKVTSHTELTSNOQSANKTQATAKOPINR 1920  
 Db 1861 SLSSLDFFDDDDVLSREKAEIRAKENKESFAKVTSHTELTSNOQSANKTQATAKOPINR 1920  
 QY 1921 GQPKPILQKOSTFQSSKIDPDRGAATDEKLQNAFIENTPVCFSHNSLSLSDIDQENN 1980  
 Db 1921 GQPKPILQKOSTFQSSKIDPDRGAATDEKLQNAFIENTPVCFSHNSLSLSDIDQENN 1980  
 QY 1981 NKNEPIKETEPDPSQOEPKQASGAPKSFHVEDTPVCFSRNSLSLSIDSEDDLLQ 2040  
 Db 1981 NKNEPIKETEPDPSQOEPKQASGAPKSFHVEDTPVCFSRNSLSLSIDSEDDLLQ 2040  
 QY 2041 ECISAMPKPKKRLKGDNEKHSFRNMGILGEDTLDLKDIOQPDSEHGLSPDSENF 2100  
 Db 2041 ECISAMPKPKKRLKGDNEKHSFRNMGILGEDTLDLKDIOQPDSEHGLSPDSENF 2100  
 QY 2101 WKATQEGANSIVSSLHQAANAALSRQASSDSLSLSKSGISLGSFHLTPDQEEKPFT 2160  
 Db 2101 WKATQEGANSIVSSLHQAANAALSRQASSDSLSLSKSGISLGSFHLTPDQEEKPFT 2160  
 QY 2161 SNKGPRILKPEKSTLTETKTESKGIKGGKVKYKSLITGKVRNSEISGQWKQLOAN 2220  
 Db 2161 SNKGPRILKPEKSTLTETKTESKGIKGGKVKYKSLITGKVRNSEISGQWKQLOAN 2220  
 QY 2221 MPISIRGRMTIHIPGVNRNSSSTSPVSKGPPPLKTPASKSPSEGQTATTSPRGAKPSVKS 2280  
 Db 2221 MPISIRGRMTIHIPGVNRNSSSTSPVSKGPPPLKTPASKSPSEGQTATTSPRGAKPSVKS 2280

QY 2281 ELSPVARQTSOIGSSSKAPSRSGSRDSTPSRPAQOPLSRPIQSPGRNSISPGRNGISPPN 2340  
 Db 2281 ELSPVARQTSOIGSSSKAPSRSGSRDSTPSRPAQOPLSRPIQSPGRNSISPGRNGISPPN 2340  
 QY 2341 KLSQLPRTSSSTASTKSSGSKMSYTPGQMSQOQLTKOTGLSKNASSITPSESASKG 2400  
 Db 2341 KLSQLPRTSSSTASTKSSGSKMSYTPGQMSQOQLTKOTGLSKNASSITPSESASKG 2400  
 QY 2401 LQNMNNGANGANKVELSEMSSTKSSGSDSRSERPVLVRQSTFIKEAPSPTLRKLBEA 2460  
 Db 2401 LQNMNNGANGANKVELSEMSSTKSSGSDSRSERPVLVRQSTFIKEAPSPTLRKLBEA 2460  
 QY 2461 SFESLSPSRSPASPTRSOAQTIVLSPLPMSLTHSSVQAGGWRKLPNLSPTIEYNDG 2520  
 Db 2461 SFESLSPSRSPASPTRSOAQTIVLSPLPMSLTHSSVQAGGWRKLPNLSPTIEYNDG 2520  
 QY 2521 RPAKGRHDITARGHSESPSRPLPINRSGTWKREHSHKSSSLPRVSTWRTGSSSSILSASES 2580  
 Db 2521 RPAKGRHDITARGHSESPSRPLPINRSGTWKREHSHKSSSLPRVSTWRTGSSSSILSASES 2580  
 QY 2581 SEKAKSEDEKHVNSISGTSKQKNOVSAKGTWRKI KENEFSPNTNSTQTSVSSGATNGAES 2640  
 Db 2581 SEKAKSEDEKHVNSISGTSKQKNOVSAKGTWRKI KENEFSPNTNSTQTSVSSGATNGAES 2640  
 QY 2641 KTLIYQMAPAVSKTEDVWRIEDCPINNPRSGRPTGNTPPVIDSVSEKANPNIKDSKDN 2700  
 Db 2641 KTLIYQMAPAVSKTEDVWRIEDCPINNPRSGRPTGNTPPVIDSVSEKANPNIKDSKDN 2700  
 QY 2701 QAKQVNGSGVPMRTVGLNRLNSFIQVADAPQKGTBIKPGQNNPVPVSETNESSIVERT 2760  
 Db 2701 QAKQVNGSGVPMRTVGLNRLNSFIQVADAPQKGTBIKPGQNNPVPVSETNESSIVERT 2760  
 QY 2761 PFSSSSSSKHSPSGTVAARVTTPNNYNSPEKSSADSTSRPSOIPTPVNNNTKKRDSKT 2820  
 Db 2761 PFSSSSSSKHSPSGTVAARVTTPNNYNSPEKSSADSTSRPSOIPTPVNNNTKKRDSKT 2820  
 QY 2821 DSTESSGTQSPKSHSGSYLVTSV 2843  
 Db 2821 DSTESSGTQSPKSHSGSYLVTSV 2843  
 RESULT 12  
 AAR26052  
 ID AAR26052 standard; protein; 2843 AA.  
 XX AC AAR26052;  
 XX DT 25-MAR-2003 (revised)  
 XX DT 28-JAN-1993 (first entry)  
 XX DE APC gene product in familial adenomatous polyposis.  
 XX KW neoplasm; cancer; oncogene; tumour; growth; detection; diagnosis;  
 KW prognosis; treatment; sporadic colorectal carcinomas; ss.  
 XX OS Homo sapiens.  
 XX PN W09213103-A1.  
 XX PD 06-AUG-1992.  
 XX PF 16-JAN-1992; 92WO-US0000376.  
 XX PR 16-JAN-1991; 91CB-00000963.  
 PR 08-AUG-1991; 91US-00741940.  
 XX PA (UKJO ) UNIV JOHNS HOPKINS.  
 PA (ICIL ) IMPERIAL CHEM IND PLC.  
 PA (UTAH ) UNIV UTAH.  
 PA (CANC-) CANCER INST.  
 XX PI Kinzler KW, Vogelstein B, Anand R, Hedge PJ, Markham AF;  
 PI Albertsen H, Carlson ML, Groden JL, Joslyn G, Thliveris A, White RL;



Db 1801 AERFSDNKKONLKNVNSKDNKNDKLPNNEDVRGSPAFDSPPHHYPTIEGTPYCFSRND 1860  
 Qy 1861 SLSLDPPDDVLSRKAELRKAENKESAKVTSHTELTNSQOQANKTQAIKQPINR 1920  
 Db 1861 SLSLDPPDDVLSRKAELRKAENKESAKVTSHTELTNSQOQANKTQAIKQPINR 1920  
 Qy 1921 GQPKPILOKOSTPPOSSKDIPIRGAATDEKQNFATENTPVCFSHNSLSLSDIDQENN 1980  
 Db 1921 GQPKPILOKOSTPPOSSKDIPIRGAATDEKQNFATENTPVCFSHNSLSLSDIDQENN 1980  
 Qy 1981 NKNEPIKETEPDPOSSKQEPKQASGYAPKSFHVEPTPVCFSNSLSLSIDSEDDLIQ 2040  
 Db 1981 NKNEPIKETEPDPOSSKQEPKQASGYAPKSFHVEPTPVCFSNSLSLSIDSEDDLIQ 2040  
 Qy 2041 ECISAMPKPKKPSRLKGDNEKSPRNWGGILGEDLTLDKDIQRDPDSEHGLSPDSENF 2100  
 Db 2041 ECISAMPKPKKPSRLKGDNEKSPRNWGGILGEDLTLDKDIQRDPDSEHGLSPDSENF 2100  
 Qy 2101 WKAIQEGANSIVSSLHQAACLSROASDSDSIILSKSGISLGSPPFHLTPDQEKPT 2160  
 Db 2101 WKAIQEGANSIVSSLHQAACLSROASDSDSIILSKSGISLGSPPFHLTPDQEKPT 2160  
 Qy 2161 SNKGPRILKGEKSTLTETKIEBESKIKGGKVKYKSLITGKVRNSSEISGQWKQLOAN 2220  
 Db 2161 SNKGPRILKGEKSTLTETKIEBESKIKGGKVKYKSLITGKVRNSSEISGQWKQLOAN 2220  
 Qy 2221 MPSISRGRTHIIPGVNSSSSTSPVSKGPPPLKTPASKSPSEGQTATTPRGAKSPVKS 2280  
 Db 2221 MPSISRGRTHIIPGVNSSSSTSPVSKGPPPLKTPASKSPSEGQTATTPRGAKSPVKS 2280  
 Qy 2281 ELSVARTQTOIGSSKAPSRGSRDTPSRPAQOPLSRPIQSPGRNSISPGRNGISPPN 2340  
 Db 2281 ELSVARTQTOIGSSKAPSRGSRDTPSRPAQOPLSRPIQSPGRNSISPGRNGISPPN 2340  
 Qy 2341 KLSQLPRTSPSTASTKSSGKMSYTPSGRQSQNLTKQTLGSKNAGSIPRSEASKG 2400  
 Db 2341 KLSQLPRTSPSTASTKSSGKMSYTPSGRQSQNLTKQTLGSKNAGSIPRSEASKG 2400  
 Qy 2401 LNMWNGANGANKVELSRMSTKSSGESDRSRPVLVROSTFIKAPSPFLRRKLEESA 2460  
 Db 2401 LNMWNGANGANKVELSRMSTKSSGESDRSRPVLVROSTFIKAPSPFLRRKLEESA 2460  
 Qy 2461 SFESLSPSSRPASTRQAOTPVLSPLDMSLSTHSSVOAGGWRKLPNLSPTIENDG 2520  
 Db 2461 SFESLSPSSRPASTRQAOTPVLSPLDMSLSTHSSVOAGGWRKLPNLSPTIENDG 2520  
 Qy 2521 RPAKRHDIAHSHSPSRPLNRSGTWKREHSHSSSLPRVSTWRTGTSSSILSASSES 2580  
 Db 2521 RPAKRHDIAHSHSPSRPLNRSGTWKREHSHSSSLPRVSTWRTGTSSSILSASSES 2580  
 Qy 2581 SEKAKSEDEKHVNSISGTQSKENQVSAKGTWKIKENEPSTNSQTVSSGATNGAES 2640  
 Db 2581 SEKAKSEDEKHVNSISGTQSKENQVSAKGTWKIKENEPSTNSQTVSSGATNGAES 2640  
 Qy 2641 KTLIYQWAPAVSKTEDEVVWRIEDCPINPRSGRSPGTNTPPVIDSVSEKANPNIKDKDN 2700  
 Db 2641 KTLIYQWAPAVSKTEDEVVWRIEDCPINPRSGRSPGTNTPPVIDSVSEKANPNIKDKDN 2700  
 Qy 2701 QAKQNVGNGVPMRTVGLNRLNSFIQVADPDKGTEIKPGQNNPVPVSTNETSSIVERT 2760  
 Db 2701 QAKQNVGNGVPMRTVGLNRLNSFIQVADPDKGTEIKPGQNNPVPVSTNETSSIVERT 2760  
 Qy 2761 PFSSSSSKKSSPSGCTVAARVTPPNYNPSPRKSADSTSPRSQIPTPVNNNTKKRDSKT 2820  
 Db 2761 PFSSSSSKKSSPSGCTVAARVTPPNYNPSPRKSADSTSPRSQIPTPVNNNTKKRDSKT 2820  
 Qy 2821 DSTESSGTQSPKRHSGSYLTVS 2843  
 Db 2821 DSTESSGTQSPKRHSGSYLTVS 2843

ID AAW35392 standard; protein; 2843 AA.  
 XX  
 AC AAW35392;  
 XX  
 DT 25-MAR-2003 (revised)  
 DT 11-MAR-1998 (first entry)  
 XX  
 DE Human adenomatous Polyposis coli gene product.  
 XX  
 KW Human; adenomatous Polyposis coli; APC; diagnosis; prognosis;  
 KW neoplastic tissue; tumour tissue; tumour repressor; mutation;  
 KW sporadic colorectal cancer; detection.  
 XX  
 OS Homo sapiens.  
 XX  
 PN US5648212-A.  
 XX  
 PD 15-JUL-1997.  
 XX  
 PF 12-AUG-1994; 94US-00289548.  
 XX  
 PR 16-JAN-1991; 91GB-00000962.  
 PR 16-JAN-1991; 91GB-00000963.  
 PR 16-JAN-1991; 91GB-00000974.  
 PR 16-JAN-1991; 91GB-00000975.  
 PR 08-AUG-1991; 91US-00741940.  
 XX  
 PA (NICA-) JAPANESE FOUND CANCER RES.  
 PA (UTAH) UNIV UTAH.  
 PA (UIJO) UNIV JOHNS HOPKINS.  
 PA (ZENE) ZENECA LTD.  
 XX  
 PI Markham A, Nakamura Y, Groden J, Carlson M, Kinzler K;  
 PI Albertsen H, Hedge PJ, Vogelstein B, Thiliveris A, Anand R, White RL;  
 PI Joslyn G;  
 XX  
 DR WPI: 1997-372053/34.  
 DR N-PSDB; AAR95538.  
 XX  
 PT Cancer diagnosis - by detecting mutation(s) in adenomatous polyposis coli  
 PT gene.  
 XX  
 PS Example 1; Col 33-52; 140pp; English.  
 XX  
 CC The present sequence is the human adenomatous Polyposis coli (APC) gene  
 CC product, which was used in the development of a novel method of  
 CC diagnosing or prognosing an APC gene associated neoplastic tissue. The  
 CC method comprises comparing APC gene coding sequences or mRNA in a tumour  
 CC tissue, to APC gene coding sequences or mRNA in a non-neoplastic tissue,  
 CC where a difference indicates an APC gene associated neoplasia of the  
 CC tumour tissue. APC is a tumour repressor expressed in most normal  
 CC tissues. APC mutations are found in familial adenomatous polyposis and  
 CC sporadic colorectal cancer patients. The method enables mutations to be  
 CC detected to provide an indication of predisposition to cancer. (Updated  
 CC on 25-MAR-2003 to correct PR field.)  
 XX  
 SQ Sequence 2843 AA;

Query Match 99.8%; Score 14533; DB 2; Length 2843;  
 Best Local Similarity 99.8%; Pred. No. 0;  
 Matches 2836; Conservative 2; Mismatches 5; Indels 0; Gaps 0;  
 Qy 1 MAASDYDOLLKQVEALKMENSNLQLEEDNSNHLTKLETEASNKKEVLKQIQGSIEDEAM 60  
 Db 1 MAASDYDOLLKQVEALKMENSNLQLEEDNSNHLTKLETEASNKKEVLKQIQGSIEDEAM 60  
 Qy 61 ASSGOIDLLERLKEINLDSNFPQVKLRKMSLSYSGSRGSSVSRGSGSPVPMGSPFR 120  
 Db 61 ASSGOIDLLERLKEINLDSNFPQVKLRKMSLSYSGSRGSSVSRGSGSPVPMGSPFR 120  
 Qy 121 RGFVNGSRSTGYLEELERKERSLLIADLDKEEKDWYQALQNLTKRIDSPLTENFSL 180  
 Db 121 RGFVNGSRSTGYLEELERKERSLLIADLDKEEKDWYQALQNLTKRIDSPLTENFSL 180

181 QY QDMDTRRQLEVEARQIRVAMEEQLGTCQDMKRAQRRIARIQOIEKQILRIRQLLQSOAT 240  
 181 DB OTDLTRRQLEVEARQIRVAMEEQLGTCQDMKRAQRRIARIQOIEKQILRIRQLLQSOAT 240  
 241 QY BAERSQNKHETGSDAERONEGQGVGINNATSGNGOGSTTRMDHETASVLSSSSTHSA 300  
 241 DB BAERSQNKHETGSDAERONEGQGVGINNATSGNGOGSTTRMDHETASVLSSSSTHSA 300  
 301 QY PRLTSHLGTQVEMVYSLMLGTHDKDDMRTLLAMSSQDSCISMRQSCCLPILLIQLL 360  
 301 DB PRLTSHLGTQVEMVYSLMLGTHDKDDMRTLLAMSSQDSCISMRQSCCLPILLIQLL 360  
 361 QY HGNKDXSVLLGNSRGSKAPARASAAALNNIHSQDDXGRBRETRVLHLLQIRAYCETC 420  
 361 DB HGNKDXSVLLGNSRGSKAPARASAAALNNIHSQDDXGRBRETRVLHLLQIRAYCETC 420  
 421 QY WEQEAHEPFGDODKNPMPAPVEHQICPAVCVLMKLSFDEEHRHAMNELGLQIAELQ 480  
 421 DB WEQEAHEPFGDODKNPMPAPVEHQICPAVCVLMKLSFDEEHRHAMNELGLQIAELQ 480  
 481 QY VDCENYGLTNDHYSITLRYAGMALTNITFGDVANKATLCNKGCMRALVAQLKSEBSDL 540  
 481 DB VDCENYGLTNDHYSITLRYAGMALTNITFGDVANKATLCNKGCMRALVAQLKSEBSDL 540  
 541 QY QOVIASVLRLNLSWRADVNSKTLREVGSVKALMECALEVKXESTLKSVLKALWNLASHCT 600  
 541 DB QOVIASVLRLNLSWRADVNSKTLREVGSVKALMECALEVKXESTLKSVLKALWNLASHCT 600  
 601 QY ENKADICAVDGAFLVGLTLYRSQNTNLAIIESGGGLIRNVSSLIATNEDHROILRENN 660  
 601 DB ENKADICAVDGAFLVGLTLYRSQNTNLAIIESGGGLIRNVSSLIATNEDHROILRENN 660  
 661 QY CLQTLLOHLKSHSLTIVSNACQTLNLSARPKDQOALWDMGAVSMKLNLIHSKXKMIAM 720  
 661 DB CLQTLLOHLKSHSLTIVSNACQTLNLSARPKDQOALWDMGAVSMKLNLIHSKXKMIAM 720  
 721 QY GSAALRLNMANRPKAKYKADANIMSGSSLPSSHVRKQKALAEALDAQHLSETFONIDNLS 780  
 721 DB GSAALRLNMANRPKAKYKADANIMSGSSLPSSHVRKQKALAEALDAQHLSETFONIDNLS 780  
 781 QY PKASHRSQRKHQSLYGDYVFTNDRHDDNRSNFTNMTNVLSPYLNTTVLPSSSSSRGS 840  
 781 DB PKASHRSQRKHQSLYGDYVFTNDRHDDNRSNFTNMTNVLSPYLNTTVLPSSSSSRGS 840  
 841 QY LQSSRSSEKDRSLERERGLGNYPHATENPGTSSKRGLOISTTAAQIAKWEEVSAIHTS 900  
 841 DB LQSSRSSEKDRSLERERGLGNYPHATENPGTSSKRGLOISTTAAQIAKWEEVSAIHTS 900  
 901 QY QEDRSSGSTTELHCVTDERNALRRSSAAHTSNFTKXSENSNRTCSEMPYAKLEYKRSS 960  
 901 DB QEDRSSGSTTELHCVTDERNALRRSSAAHTSNFTKXSENSNRTCSEMPYAKLEYKRSS 960  
 961 QY NQSLNSVSSSDGYKRGQWKPSIESYEDDESFKCSYQYPADIAHKTHSANHMDNDGE 1020  
 961 DB NQSLNSVSSSDGYKRGQWKPSIESYEDDESFKCSYQYPADIAHKTHSANHMDNDGE 1020  
 1021 QY LQTPINYSILKYDEQLNSGRQSPQNERWARPKHIEDEIKQEQORQSRNQSTTYPVYTE 1080  
 1021 DB LQTPINYSILKYDEQLNSGRQSPQNERWARPKHIEDEIKQEQORQSRNQSTTYPVYTE 1080  
 1081 QY STDDKHLKQPHFGQOECVPSYRSGANGSETNRVGSNHGINQVNSQSLQEDDVEDDKP 1140  
 1081 DB STDDKHLKQPHFGQOECVPSYRSGANGSETNRVGSNHGINQVNSQSLQEDDVEDDKP 1140  
 1141 QY TNYSERYSBEEQHEERPTNYSIKYNEEKRVHVDOPIDYSLKYATDIPSSOKQSPFSKS 1200  
 1141 DB TNYSERYSBEEQHEERPTNYSIKYNEEKRVHVDOPIDYSLKYATDIPSSOKQSPFSKS 1200  
 1201 QY SSGQSSKTEHMSSESSNTSTPSSNAKRONQLHPSSAQSRSGQPOKAATCKVSSINQETIQ 1260  
 1201 DB SSGQSSKTEHMSSESSNTSTPSSNAKRONQLHPSSAQSRSGQPOKAATCKVSSINQETIQ 1260

1261 QY TYCVETPTICFSSCSLSLSAEDBIGNQTTQEADSANTLQIAIBIKKIGTRSAEDPV 1320  
 1261 DB TYCVETPTICFSSCSLSLSAEDBIGNQTTQEADSANTLQIAIBIKKIGTRSAEDPV 1320  
 1321 QY SEVPANVQBPRTKSRLOQSSLSLSSESARHKAVEFFSGAKSPSKSGAQTPKSPPEHYVQET 1380  
 1321 DB SEVPANVQBPRTKSRLOQSSLSLSSESARHKAVEFFSGAKSPSKSGAQTPKSPPEHYVQET 1380  
 1381 QY PLMFSTRCTSVSSLDSPESRSIASSVQSEPCSGMVSGIISPSDLPPSGQTMPPSRSKTTP 1440  
 1381 DB PLMFSTRCTSVSSLDSPESRSIASSVQSEPCSGMVSGIISPSDLPPSGQTMPPSRSKTTP 1440  
 1441 QY PPQTAQTKKEEVPKNKAPTAEKESGPKQAAVNAVQVQVLPDADTLAHFATESPDGF 1500  
 1441 DB PPQTAQTKKEEVPKNKAPTAEKESGPKQAAVNAVQVQVLPDADTLAHFATESPDGF 1500  
 1501 QY SCSSLSALSALDEFFIQQOVLELIMPPVQENDNGNETESEQPKESNENQKEAEKTIDSE 1560  
 1501 DB SCSSLSALSALDEFFIQQOVLELIMPPVQENDNGNETESEQPKESNENQKEAEKTIDSE 1560  
 1561 QY KDLLDSDDDDDIIEBECIISAMPTKSSRKAKKPAQATASKLPPPVARKPSOLPVYKLLPS 1620  
 1561 DB KDLLDSDDDDDIIEBECIISAMPTKSSRKAKKPAQATASKLPPPVARKPSOLPVYKLLPS 1620  
 1621 QY QNRLQPKQHVSTPFGDDMPRVYCVEGTPIINFSTATLSLDLTIESPPNELAAGSGVRGGAQ 1680  
 1621 DB QNRLQPKQHVSTPFGDDMPRVYCVEGTPIINFSTATLSLDLTIESPPNELAAGSGVRGGAQ 1680  
 1681 QY SGFEKEDTITPTGRSTDEAQGGKTSSVTIPELDDNKAEBGDIILAEICINSAMPKGSKHP 1740  
 1681 DB SGFEKEDTITPTGRSTDEAQGGKTSSVTIPELDDNKAEBGDIILAEICINSAMPKGSKHP 1740  
 1741 QY FRVKKIMQVQOQASASSAPNKNQLOGKKKPTSPVKPIPONTEYTRVRKNADSKNNLN 1800  
 1741 DB FRVKKIMQVQOQASASSAPNKNQLOGKKKPTSPVKPIPONTEYTRVRKNADSKNNLN 1800  
 1801 QY AERVFSNDKSKQNLKNSKDFNDKLPNNEDVRGSAFDPSPHYTPIEGTGYCFSRND 1860  
 1801 DB AERVFSNDKSKQNLKNSKDFNDKLPNNEDVRGSAFDPSPHYTPIEGTGYCFSRND 1860  
 1861 QY SLASLDEDDDDVLSREKAEELKAKENKSEAKVTSHTELTNQOQANKTQAAKAPINR 1920  
 1861 DB SLASLDEDDDDVLSREKAEELKAKENKSEAKVTSHTELTNQOQANKTQAAKAPINR 1920  
 1921 QY GQKPILOKQSTFPQSSKDIIPDRGAATDEKLNQFAIENTPVCFSHNSLSLSDIDQENN 1980  
 1921 DB GQKPILOKQSTFPQSSKDIIPDRGAATDEKLNQFAIENTPVCFSHNSLSLSDIDQENN 1980  
 1981 QY NKENEPKETEPPDQSGEPKQASGYAPKSFHVEDTPVCFSRNSLSLSLSDIDQENN 2040  
 1981 DB NKENEPKETEPPDQSGEPKQASGYAPKSFHVEDTPVCFSRNSLSLSLSDIDQENN 2040  
 2041 QY ECISAMPKPKKPSRLKGDNEKHSPRNMGGILGEDTLDLKDIQRPDSEHGLSPDSENF 2100  
 2041 DB ECISAMPKPKKPSRLKGDNEKHSPRNMGGILGEDTLDLKDIQRPDSEHGLSPDSENF 2100  
 2101 QY WKAIQEGANSIVSLHQAACCLSRQASSDSDSILSLKSGISLGSFPHLTPOQEEKPFT 2160  
 2101 DB WKAIQEGANSIVSLHQAACCLSRQASSDSDSILSLKSGISLGSFPHLTPOQEEKPFT 2160  
 2161 QY SNKGPRILKPEKXSTLETKKIETESKGIKGGKVKYKSLITGKVRNSSEISQOMKQLOAN 2220  
 2161 DB SNKGPRILKPEKXSTLETKKIETESKGIKGGKVKYKSLITGKVRNSSEISQOMKQLOAN 2220  
 2221 QY MPSISRGRITMIHIPGVNNSSTSPVSKGPPKLTTPASKSPSEGTATTTPRGAKPSVK 2280  
 2221 DB MPSISRGRITMIHIPGVNNSSTSPVSKGPPKLTTPASKSPSEGTATTTPRGAKPSVK 2280  
 2281 QY ELSVPARQTSIQGSSKAPSRSGRSDTPSRPAQOPLSRPIQSPGRNSISPGANGISPPN 2340  
 2281 DB ELSVPARQTSIQGSSKAPSRSGRSDTPSRPAQOPLSRPIQSPGRNSISPGANGISPPN 2340  
 2341 QY KLSQLPRTSPSTASTKSSGSGKMSYTPSGRQMSQQLTKQTGLSKNASGTPRSESASKG 2400



Db 2341 KLSQPTSTSPSTASTKSSGSKMSYTPGRQMSQQNLTKQTGLSKNASSIPSEASKG 2400  
 Qy 2401 LNMNNGNGANKVELSRMSTKSSGSDRSPVLVRQSTTIKEAPSTLRKLEESA 2460  
 Db 2401 LNMNNGNGANKVELSRMSTKSSGSDRSPVLVRQSTTIKEAPSTLRKLEESA 2460  
 Qy 2461 SFESLSPSPSPSTRSQATPVLSPLPDMKSLSTHSSVQAGGWRKLPNLSITIIYNDG 2520  
 Db 2461 SFESLSPSPSPSTRSQATPVLSPLPDMKSLSTHSSVQAGGWRKLPNLSITIIYNDG 2520  
 Qy 2521 RPAKRDHDIARSHSPRLPINRSGTWKREHSHSSSLPRVSTWRRTGSSSSITLSASSES 2580  
 Db 2521 RPAKRDHDIARSHSPRLPINRSGTWKREHSHSSSLPRVSTWRRTGSSSSITLSASSES 2580  
 Qy 2581 SEKAKSEDEKHVNSISGTQSKENQVSAKGTWRKIKENEPSPNTNSTQTVSSGATNGAES 2640  
 Db 2581 SEKAKSEDEKHVNSISGTQSKENQVSAKGTWRKIKENEPSPNTNSTQTVSSGATNGAES 2640  
 Qy 2641 KTLIYQMAPAVSKTEDVWRIEDCPINPRSGRSPGTGTPPVVDSVSEKANPNIKDSKDN 2700  
 Db 2641 KTLIYQMAPAVSKTEDVWRIEDCPINPRSGRSPGTGTPPVVDSVSEKANPNIKDSKDN 2700  
 Qy 2701 QAKQNVGSGVPMRTVGLNRLNLSFIQVDAPDQKTEIKPGQNNPVVPSVETNESSIVERT 2760  
 Db 2701 QAKQNVGSGVPMRTVGLNRLNLSFIQVDAPDQKTEIKPGQNNPVVPSVETNESSIVERT 2760  
 Qy 2761 PFSSSSSKHSSPSGTVAAAVTPFNYPNPSPRKSADSTSRPSQIPTPVNNNTKKEDSKT 2820  
 Db 2761 PFSSSSSKHSSPSGTVAAAVTPFNYPNPSPRKSADSTSRPSQIPTPVNNNTKKEDSKT 2820  
 Qy 2821 DSTESSGTQSPKRHSGSYLTVS 2843  
 Db 2821 DSTESSGTQSPKRHSGSYLTVS 2843

RESULT 14  
 AAW38370  
 ID AAW38370 standard; protein; 2843 AA.  
 XX AC AAW38370;  
 XX DT 25-MAR-2003 (revised)  
 XX DT 08-APR-1998 (first entry)  
 XX DE Human adenomatous Polyposis coli gene product.  
 XX KW Human; adenomatous Polyposis coli; APC; diagnosis; prognosis;  
 KW neoplastic tissue; tumour tissue; tumour repressor; mutation;  
 KW sporadic colorectal cancer; detection.  
 XX OS Homo sapiens.  
 XX PN US5691454-A.  
 XX PD 25-NOV-1997.  
 XX PF 25-MAY-1995; 95US-00452654.  
 XX PR 16-JAN-1991; 91GB-00000962.  
 XX PR 16-JAN-1991; 91GB-00000963.  
 XX PR 16-JAN-1991; 91GB-00000974.  
 XX PR 16-JAN-1991; 91GB-00000975.  
 XX PR 08-AUG-1991; 91US-00741940.  
 XX PR 12-AUG-1994; 94US-00289548.  
 XX PA (CANC-) CANCER INST.  
 XX PA (UJO) UNIV JOHNS HOPKINS.  
 XX PA (ICIL) IMPERIAL CHEM IND PLC.  
 XX PA (UTAH) UNIV UTAH.  
 XX PI Nakamura Y, Markham AF, Groden J, Vogelstein B, Kinzler K,  
 PI Hedge PJ, Carlson M, Anand R, Thliveris A, Albertsen H, White RL;

PI Joslyn G;  
 DR WPI; 1998-017712/02.  
 DR N-PSDB; AAT96153.  
 XX  
 PT Antibodies to normal and mutant adenomatous polyposis coli proteins -  
 useful for detecting genetic predisposition to cancer.  
 XX  
 XX Example 1; Col 33-52; 107pp; English.  
 XX  
 CC The present sequence is the human adenomatous Polyposis coli (APC) gene  
 product, which was used in the development of a novel method of  
 diagnosing or prognosing an APC gene associated neoplastic tissue. The  
 method comprises comparing APC gene coding sequences or mRNA in a tumour  
 tissue, to APC gene coding sequences or mRNA in a non-neoplastic tissue,  
 where a difference indicates an APC gene associated neoplasia of the  
 tumour tissue. APC is a tumour repressor expressed in most normal  
 tissues. APC mutations are found in familial adenomatous polyposis and  
 sporadic colorectal cancer patients. The method enables mutations to be  
 detected to provide an indication of predisposition to cancer. (Updated  
 on 25-MAR-2003 to correct PR field.)  
 XX  
 SQ Sequence 2843 AA;  
 Query Match 99.8%; Score 14533; DB 2; Length 2843;  
 Best Local Similarity 99.8%; Pred. No. 0;  
 Matches 2836; Conservative 2; Mismatches 5; Indels 0; Gaps 0;  
 Qy 1 MAASDYDQLLKQVALKMENSINLROELEDNSNHLTKLETEASNKKVVKQLQGSIEDEAM 60  
 Db 1 MAASDYDQLLKQVALKMENSINLROELEDNSNHLTKLETEASNKKVVKQLQGSIEDEAM 60  
 Qy 61 ASSGQIDLLERLKEINLDSNFPVGVKLRKMSLSYSGSRGSSVSRGEGSPVPMGSPFR 120  
 Db 61 ASSGQIDLLERLKEINLDSNFPVGVKLRKMSLSYSGSRGSSVSRGEGSPVPMGSPFR 120  
 Qy 121 RGFVNGSRSTGYLEELKERSLLADLDKEEKQWYQAQLQNLTKRIDSLPTENFSL 180  
 Db 121 RGFVNGSRSTGYLEELKERSLLADLDKEEKQWYQAQLQNLTKRIDSLPTENFSL 180  
 Qy 181 QTDLTRQLEYEARQIRVAMEEQLGTQDMKRAORRIARIQIEKDIILRIQLQSQAT 240  
 Db 181 QTDLTRQLEYEARQIRVAMEEQLGTQDMKRAORRIARIQIEKDIILRIQLQSQAT 240  
 Qy 241 EARSQQNKHETGSHDAERQNEGQVGEINMATSNGCGSTTRMDHETASVLSSTHSA 300  
 Db 241 EARSQQNKHETGSHDAERQNEGQVGEINMATSNGCGSTTRMDHETASVLSSTHSA 300  
 Qy 301 PRLTSHLGTVMYVYSLIIMLGTDKDMSRTLLAMSSQDSCISMKGCCPLLIQLL 360  
 Db 301 PRLTSHLGTVMYVYSLIIMLGTDKDMSRTLLAMSSQDSCISMKGCCPLLIQLL 360  
 Qy 361 HGNDKDSVLLGNSRGSKEARASAAALHNIHSQPDQKGRREIRVHLLEQIRAYCETC 420  
 Db 361 HGNDKDSVLLGNSRGSKEARASAAALHNIHSQPDQKGRREIRVHLLEQIRAYCETC 420  
 Qy 421 WEQBAHFGMDQDNPMAPVEHQICPAVCVLMKLSFDEEHRHANNELGLOAIELLQ 480  
 Db 421 WEQBAHFGMDQDNPMAPVEHQICPAVCVLMKLSFDEEHRHANNELGLOAIELLQ 480  
 Qy 481 VDCMYGLTNDHYISITLRRVAGMALTNLTGFDVANKATLCMKGCMALVAQKSEEDL 540  
 Db 481 VDCMYGLTNDHYISITLRRVAGMALTNLTGFDVANKATLCMKGCMALVAQKSEEDL 540  
 Qy 541 QQVIASVLRNLSWRADVNSKKTLEVGSVTALMECALEVKESTLKSIVLSALNLSAHT 600  
 Db 541 QQVIASVLRNLSWRADVNSKKTLEVGSVTALMECALEVKESTLKSIVLSALNLSAHT 600  
 Qy 601 ENKADICAVDQALAFVLTITRSQNTNLAIISGGGILNRVSSLIATNEDHRIQILRENN 660  
 Db 601 ENKADICAVDQALAFVLTITRSQNTNLAIISGGGILNRVSSLIATNEDHRIQILRENN 660  
 Qy 661 CLQTLQHLKSHSLTIVSNACGLTNLWLSARNPKDQEAALWDMGAVSMUKNLIHSHKHMIA 720



661	Db		CLQTLQLKSHSLITVSNACGTLWNLARSNPDKQALMDGAVSMLKXLIHSKHOMIAM	720
721	Qy	721	GSAAALRNLMANRPAXYKDANIMSGSSILPSLHVRKOKALEAELDAQHSETFDNI.DNLS	780
721	Db	721	GSAAALRNLMANRPAXYKDANIMSGSSILPSLHVRKOKALEAELDAQHSETFDNI.DNLS	780
781	Qy	781	PKASHRSKORHKOSLYGDYVFDNRHDDNRSDNFNTGNMTVLSPYLNTTVLPSSSSSRGS	840
781	Db	781	PKASHRSKORHKOSLYGDYVFDNRHDDNRSDNFNTGNMTVLSPYLNTTVLPSSSSSRGS	840
841	Qy	841	LDSSRSEKORSLEPERGIGLGNVHPATEPGTSSKGLQISTTAAQIAKVMEEVSAIHTS	900
841	Db	841	LDSSRSEKORSLEPERGIGLGNVHPATEPGTSSKGLQISTTAAQIAKVMEEVSAIHTS	900
901	Qy	901	QEDRSSGSTEHCVTDERNALRRSSAAHTHSNTYNTKSENSNRTCSPYAKLEYKRSS	960
901	Db	901	QEDRSSGSTEHCVTDERNALRRSSAAHTHSNTYNTKSENSNRTCSPYAKLEYKRSS	960
961	Qy	961	NDSLNVSVDGCGKQOMKPSIESYSEDESXFCVGOYPADLAHKIHSANHMDNDGGE	1020
961	Db	961	NDSLNVSVDGCGKQOMKPSIESYSEDESXFCVGOYPADLAHKIHSANHMDNDGGE	1020
1021	Qy	1021	LDTPINYSKYSEQLNSGRQSPSONERWAPKPHIIEDEIKQEQQRNQSTTYPVYTE	1080
1021	Db	1021	LDTPINYSKYSEQLNSGRQSPSONERWAPKPHIIEDEIKQEQQRNQSTTYPVYTE	1080
1081	Qy	1081	STDCKILKPOPHGQOECVSPYPSRGANGSETNRVGSNHGINQNVQSILCOEDDYEDDXP	1140
1081	Db	1081	STDCKILKPOPHGQOECVSPYPSRGANGSETNRVGSNHGINQNVQSILCOEDDYEDDXP	1140
1141	Qy	1141	TNYSEYSEBEEQHEEBERPTNYIKYNEEKRRHVDQPIDYSLKYATDIPSSQKQSFQSKS	1200
1141	Db	1141	TNYSEYSEBEEQHEEBERPTNYIKYNEEKRRHVDQPIDYSLKYATDIPSSQKQSFQSKS	1200
1201	Qy	1201	SSGQSKTEHMSSESSENTTSPSNAKRONQLHPSSAQSRGQPKAATCKVSSINOETIQ	1260
1201	Db	1201	SSGQSKTEHMSSESSENTTSPSNAKRONQLHPSSAQSRGQPKAATCKVSSINOETIQ	1260
1261	Qy	1261	TYCVEDTPICFRCSLSLSSAASDEIGCNOQTQEADSAANTLOIAIEIKKIGTRSAEDPV	1320
1261	Db	1261	TYCVEDTPICFRCSLSLSSAASDEIGCNOQTQEADSAANTLOIAIEIKKIGTRSAEDPV	1320
1321	Qy	1321	SEYPAVSQHPRTKSSRLQGSLSLSSSARHKAVFSSGAKSPSKGAQTPKSPPEHVQET	1380
1321	Db	1321	SEYPAVSQHPRTKSSRLQGSLSLSSSARHKAVFSSGAKSPSKGAQTPKSPPEHVQET	1380
1381	Qy	1381	PLMFSRCTSVSLDSPESRSTIASVQSEPCSGMVSGIISPSDLDPSPGOMTPSRSKTTP	1440
1381	Db	1381	PLMFSRCTSVSLDSPESRSTIASVQSEPCSGMVSGIISPSDLDPSPGOMTPSRSKTTP	1440
1441	Qy	1441	PPPOTAQTKREVKNKAPTAKEKRESGPKQAAVNAAVQRVLPDADTLHLFAFATESPDGF	1500
1441	Db	1441	PPPOTAQTKREVKNKAPTAKEKRESGPKQAAVNAAVQRVLPDADTLHLFAFATESPDGF	1500
1501	Qy	1501	SCSSLSALSILDEFFIOKVELRIMPVQENDNGNETSEPOKSENQOSKEAKTIDSE	1560
1501	Db	1501	SCSSLSALSILDEFFIOKVELRIMPVQENDNGNETSEPOKSENQOSKEAKTIDSE	1560
1561	Qy	1561	KDILDDSDDDDDIEILEBECIISAMPTKSSRKAKPAQATSKLPPVPKPSQLPVKLLPS	1620
1561	Db	1561	KDILDDSDDDDDIEILEBECIISAMPTKSSRKAKPAQATSKLPPVPKPSQLPVKLLPS	1620
1621	Qy	1621	QNRLOPKQHSVFTPGDDMPVYCVGEGTPIPNFSTATSLSDLTIESPPNELAAGGVRGGAQ	1680
1621	Db	1621	QNRLOPKQHSVFTPGDDMPVYCVGEGTPIPNFSTATSLSDLTIESPPNELAAGGVRGGAQ	1680
1681	Qy	1681	SGFEKEDTITPTEGRSTDQAGGKTSVTIPELDDNKAEBGDILAEICINSAMPKGSHPK	1740
1681	Db	1681	SGFEKEDTITPTEGRSTDQAGGKTSVTIPELDDNKAEBGDILAEICINSAMPKGSHPK	1740
1741	Qy	1741	FRVKKINDVOQASASSAPFNQOLDGKKKPTSPVKPIPONTETRYTRVRKNADSKNLN	1800

1741	Db	FRVKIMDQVOQASASSAPNKNQLOGKKKPTSPVKPPTPQNTYRTRVRKNADSKNNLN	1801
1801	Qy	AERVFDNDKSKQNLKNWSKDFNDKLPNNEDVRGSAFDPSPHHYTPTEGTYPYCFSRND	1861
1801	Db	AERVFDNDKSKQNLKNWSKDFNDKLPNNEDVRGSAFDPSPHHYTPTEGTYPYCFSRND	1861
1861	Qy	SLSSLPDDDDVDLSREKAELRKAKENKSEAKVTSHTELTSNQOANKTCATAKOPINR	1920
1861	Db	SLSSLPDDDDVDLSREKAELRKAKENKSEAKVTSHTELTSNQOANKTCATAKOPINR	1920
1921	Qy	GQPKPILQKQSTFPQSSKIDIPDRGAATDEKLQNFAIENTPVCFSHNSLSLSDIDQENN	1980
1921	Db	GQPKPILQKQSTFPQSSKIDIPDRGAATDEKLQNFAIENTPVCFSHNSLSLSDIDQENN	1980
1981	Qy	KNENEPITETEPDPSQCEPSKPOASGYAPKSFVVEDTTPVCFSRNSLSLSIDSEDDLQ	2040
1981	Db	KNENEPITETEPDPSQCEPSKPOASGYAPKSFVVEDTTPVCFSRNSLSLSIDSEDDLQ	2040
2041	Qy	ECISSAMPKKKPSRLKGDNKEKHSPRNMGILGEDTLDLKDQRPDSEHGLSPDSENF	2100
2041	Db	ECISSAMPKKKPSRLKGDNKEKHSPRNMGILGEDTLDLKDQRPDSEHGLSPDSENF	2100
2101	Qy	WKAIQEGANSIVSLHQAAAAACLSROASDSDSILSLKSGISLGSPFHLTPQEEKPF	2160
2101	Db	WKAIQEGANSIVSLHQAAAAACLSROASDSDSILSLKSGISLGSPFHLTPQEEKPF	2160
2161	Qy	SNKGPRILKEGEKSTLETTKIIESESGIKGKKVKYKSLITGKVRNSEISGQMKQPLQAN	2220
2161	Db	SNKGPRILKEGEKSTLETTKIIESESGIKGKKVKYKSLITGKVRNSEISGQMKQPLQAN	2220
2221	Qy	MPISIRGRITWHPGVNRNSSSTSPVSKGPPILKTPASKSPSEGOATTSPRGAKPVS	2280
2221	Db	MPISIRGRITWHPGVNRNSSSTSPVSKGPPILKTPASKSPSEGOATTSPRGAKPVS	2280
2281	Qy	ELSPVARTSQIGSSKAPSRSGRSDTPRPAQPLSRPIQSPGRNSISFGRNGISPPN	2340
2281	Db	ELSPVARTSQIGSSKAPSRSGRSDTPRPAQPLSRPIQSPGRNSISFGRNGISPPN	2340
2341	Qy	KLSQLPRTSPSTASTKSSGKXMSYTPSGRQMSQQNLTKQTGLSKNASSIPRSESASK	2400
2341	Db	KLSQLPRTSPSTASTKSSGKXMSYTPSGRQMSQQNLTKQTGLSKNASSIPRSESASK	2400
2401	Qy	LNQMNNGANKKVVELSRMSTKSSGESDESRRPVLVROSTFKEAPSTLREKLHESA	2460
2401	Db	LNQMNNGANKKVVELSRMSTKSSGESDESRRPVLVROSTFKEAPSTLREKLHESA	2460
2461	Qy	SFESLSPSSRPASPTRSOAQTPVLSPSLPDMSLSTHSSVOAGGRKLPPNLSPTEYNDG	2520
2461	Db	SFESLSPSSRPASPTRSOAQTPVLSPSLPDMSLSTHSSVOAGGRKLPPNLSPTEYNDG	2520
2521	Qy	RPAKRHDIAKSHSPSRBLPINRGCTWKREHKSHSSLPVSTWRTTSGSSSILSASES	2580
2521	Db	RPAKRHDIAKSHSPSRBLPINRGCTWKREHKSHSSLPVSTWRTTSGSSSILSASES	2580
2581	Qy	SEKAKSEDEKHNVISGTTKQSKENQVSAKGTWRKIKENEFSPNTSTQTVSSGATNGAES	2640
2581	Db	SEKAKSEDEKHNVISGTTKQSKENQVSAKGTWRKIKENEFSPNTSTQTVSSGATNGAES	2640
2641	Qy	KTLIYQMAPAVSKTDETVVWRIEDCFPINNPSGRSPGTGNTPPVDSVEKAMPNIKOSKDN	2700
2641	Db	KTLIYQMAPAVSKTDETVVWRIEDCFPINNPSGRSPGTGNTPPVDSVEKAMPNIKOSKDN	2700
2701	Qy	QAKQNVNGSPVMEVTGULENRLNFSFIOVDAPDQKGTIKPGQNNPVPVSEINBESSIVERT	2760
2701	Db	QAKQNVNGSPVMEVTGULENRLNFSFIOVDAPDQKGTIKPGQNNPVPVSEINBESSIVERT	2760
2761	Qy	PFSSSSSSKSHSPSGTVAARVTPPNYNPSPRKSADSTSPRPSQIPTPVNNNTKKRDSKT	2820
2761	Db	PFSSSSSSKSHSPSGTVAARVTPPNYNPSPRKSADSTSPRPSQIPTPVNNNTKKRDSKT	2820
2821	Qy	DSTESSGTSQSPKRHSGSYLVTSV	2843
2821	Db	DSTESSGTSQSPKRHSGSYLVTSV	2843

RESULT 15  
 ID ABG90964 standard; protein; 2843 AA.  
 XX ABG90964;  
 AC ABG90964;  
 DT 29-NOV-2002 (first entry)  
 XX Human adenomatous polyposis coli (APC) protein #1.  
 DE Adenomatous polyposis coli; APC; human; neoplastic tissue;  
 KW mutation detection; tumour; cancer.  
 XX Homo sapiens.  
 OS US6413727-B1.  
 PN 02-JUL-2002.  
 PD 25-MAY-1995; 95US-00449731.  
 XX 16-JAN-1991; 91GB-00000962.  
 PR 16-JAN-1991; 91GB-00000963.  
 PR 16-JAN-1991; 91GB-00000974.  
 PR 16-JAN-1991; 91GB-00000975.  
 PR 08-AUG-1991; 91US-00741940.  
 PR 12-AUG-1994; 94US-00289548.  
 XX (UYJO ) UNIV JOHNS HOPKINS.  
 PA (UTAH ) UNIV UTAH.  
 PA (NICA-) JAPANESE FOUND CANCER RES.  
 PA (ZENE ) ZENECA LTD.  
 XX Albertsen H, Anand R, Carlson M, Groden J, Hedge PU, Joslyn G;  
 PI Kinzler K, Markham AF, Nakamura Y, Thliveris A, Vogelstein B;  
 PI White RL;  
 XX WPI: 2002-641559/69.  
 DR N-PSDB; ABS67119.  
 DR Method to aid in the diagnosis/prognosis of neoplastic tissues in humans,  
 PT by detecting somatic alteration of wild-type APC protein in tumor tissue  
 PT isolated from human, the alteration indicating neoplasia of the tissue.  
 XX Disclosure; Fig 3; 140pp; English.  
 PS This invention relates to a novel method to aid in the diagnosis or  
 CC prognosis of a neoplastic tissue of a human. The method involves  
 CC detecting somatic alteration of wild-type adenomatous polyposis coli)  
 CC protein in a tumour tissue isolated from a human (the alteration  
 CC indicating neoplasia of the tissue). The method of the invention is  
 CC useful in diagnosis or prognosis of a neoplastic tissue of a human. the  
 CC method is useful in detection of genetic predisposition to cancer. The  
 CC present sequence represents a protein sequence used in the method of the  
 CC invention  
 XX Sequence 2843 AA;  
 Query Match 99.8%; Score 14533; DB 5; Length 2843;  
 Best Local Similarity 99.8%; Pred. NO. 0;  
 Matches 2836; Conservative 2; Mismatches 5; Indels 0; Gaps 0;  
 QY 1 MAAASYDQLLKQVEALKMENSNLRLQELDNSNHLTKLETEASNKVEVLKQLQGSIEDEAM 60  
 DB 1 MAAASYDQLLKQVEALKMENSNLRLQELDNSNHLTKLETEASNKVEVLKQLQGSIEDEAM 60  
 QY 61 ASSGQIDLLERLKEINLDSNFFGVKLRKMSLRYSYSGSRSRSGSCSPVPVGSPPR 120  
 DB 61 ASSGQIDLLERLKEINLDSNFFGVKLRKMSLRYSYSGSRSRSGSCSPVPVGSPPR 120  
 QY 121 RGFVNGSRSTGYLEELEKERSLLADLDKEEKEKDWYQAQLNLTKRIDSPLTENFSL 180

DB 121 RGFVNGSRSTGYLEELEKERSLLADLDKEEKEKDWYQAQLNLTKRIDSPLTENFSL 180  
 QY 181 CTDMTRRQLEYEAREQIRVAMEEQIGTODMEKRAQRRIARIQQIEKDIRLRIQLQSQAT 240  
 DB 181 CTDMTRRQLEYEAREQIRVAMEEQIGTODMEKRAQRRIARIQQIEKDIRLRIQLQSQAT 240  
 QY 241 EAERSSQNKHETGSHDAERQNEGQGVGEINMATSGNGQSTTRMDHETASVLSSSSTHSA 300  
 DB 241 EAERSSQNKHETGSHDAERQNEGQGVGEINMATSGNGQSTTRMDHETASVLSSSSTHSA 300  
 QY 301 PRLTSHLGTKEVMYVYSLLSMLGTHDKDDMSRTLLAMSSQSDSCISMRQSGCLPLLIQLL 360  
 DB 301 PRLTSHLGTKEVMYVYSLLSMLGTHDKDDMSRTLLAMSSQSDSCISMRQSGCLPLLIQLL 360  
 QY 361 HGNDKDSVLLGNRSRGSKEARARASAAALHNIHSOPDDKRGREIRVLHLLLEQIRAYCETC 420  
 DB 361 HGNDKDSVLLGNRSRGSKEARARASAAALHNIHSOPDDKRGREIRVLHLLLEQIRAYCETC 420  
 QY 421 WEMQEAHEPGMDQDNMPAPVEHQICPAVCVLMKLSFDEEHRHAMNELGGLQIAIELLQ 480  
 DB 421 WEMQEAHEPGMDQDNMPAPVEHQICPAVCVLMKLSFDEEHRHAMNELGGLQIAIELLQ 480  
 QY 481 VDCMYGLTNDHYSTILRRYAGVATLNLTFGDVANKATLCMKGCMRALVAQLKSEEDL 540  
 DB 481 VDCMYGLTNDHYSTILRRYAGVATLNLTFGDVANKATLCMKGCMRALVAQLKSEEDL 540  
 QY 541 QCVIASVLRNLSWRADVNSKTLREVGSVKALMECALEVKKESTLKSVALNLSAHCT 600  
 DB 541 QCVIASVLRNLSWRADVNSKTLREVGSVKALMECALEVKKESTLKSVALNLSAHCT 600  
 QY 601 ENKADICAVDGCALAFNCTLTYSOTNTLAIIESGGGILRNVSLLIATNEDHRLRENN 660  
 DB 601 ENKADICAVDGCALAFNCTLTYSOTNTLAIIESGGGILRNVSLLIATNEDHRLRENN 660  
 QY 661 CLQTLQHLKSHSLTIVSNACGTLNLSARNPKDQEAALWDMGAVSMKLNLIHSHKHMIAM 720  
 DB 661 CLQTLQHLKSHSLTIVSNACGTLNLSARNPKDQEAALWDMGAVSMKLNLIHSHKHMIAM 720  
 QY 721 GSAAALRNLMANRPAPKYDANIMSPGSSLSLHVYKOKALEAEALDAHLSETFDNIDLS 780  
 DB 721 GSAAALRNLMANRPAPKYDANIMSPGSSLSLHVYKOKALEAEALDAHLSETFDNIDLS 780  
 QY 781 PKASHRSKQRHKQSLYGDYVFDTNRHDDNSDNFTGNMTVLSPLYLNTTVLPSSSSSRGS 840  
 DB 781 PKASHRSKQRHKQSLYGDYVFDTNRHDDNSDNFTGNMTVLSPLYLNTTVLPSSSSSRGS 840  
 QY 841 LDSRSRSDSLERERIGLGNVHPATENPGTSKRGLOISTTAAQIAKMEVEVSAIHTS 900  
 DB 841 LDSRSRSDSLERERIGLGNVHPATENPGTSKRGLOISTTAAQIAKMEVEVSAIHTS 900  
 QY 901 QEDRSSSTTELCHVTDERNALRPSAAHNSHTNFTKSENSNRTCSMPYAKLEYKRSS 960  
 DB 901 QEDRSSSTTELCHVTDERNALRPSAAHNSHTNFTKSENSNRTCSMPYAKLEYKRSS 960  
 QY 961 NDSLNSVSSSDGYGKRGOMKPSIESYSEDESKPCSYGYPADLAHKAHNSHMDNDGE 1020  
 DB 961 NDSLNSVSSSDGYGKRGOMKPSIESYSEDESKPCSYGYPADLAHKAHNSHMDNDGE 1020  
 QY 1021 LDTPIYSLKYSDQALNSGRSQSPQNERWARPKHITEDIKQSEQRNSQSTTYPVYTE 1080  
 DB 1021 LDTPIYSLKYSDQALNSGRSQSPQNERWARPKHITEDIKQSEQRNSQSTTYPVYTE 1080  
 QY 1081 STDDKHLKFPFHQQECVSPYRSRGANGSETNRVGNHGINQNVSLQCEDDYEDDKP 1140  
 DB 1081 STDDKHLKFPFHQQECVSPYRSRGANGSETNRVGNHGINQNVSLQCEDDYEDDKP 1140  
 QY 1141 TNYSERVSEEEHHEERPTNYSIKYNEKRVHDOPIDYSLKYATDIPSSOKSEFSKS 1200  
 DB 1141 TNYSERVSEEEHHEERPTNYSIKYNEKRVHDOPIDYSLKYATDIPSSOKSEFSKS 1200  
 QY 1201 SSGOSSKTEHMSSESSENTTSSNAKQNLHPSSAQRSGQPQKAATCKVSSINQETIQ 1260

Db 1201 SSGQSKTEHSSSENTSTFSSNAKRONQLHPSSAQSRSGQPQKAATCKVSSINQETIQ 1260  
QY 1261 TYCVEDTPICFRCSGLSSAEDIEICNTOBADSANTLOIAEIKEKIGTSAEDPV 1320  
Db 1261 TYCVEDTPICFRCSGLSSAEDIEICNTOBADSANTLOIAEIKEKIGTSAEDPV 1320  
QY 1321 SEVPASQHPRTKSRRLQCSLSSSARHKAVERSSGAKSPKSGAOTPKSPPHYVQET 1380  
Db 1321 SEVPASQHPRTKSRRLQCSLSSSARHKAVERSSGAKSPKSGAOTPKSPPHYVQET 1380  
QY 1381 PLMFRCSTVSLSDFERSSTASSVQSPCGSMVSGIISPSDLDPSPGOTMPPRSKTPP 1440  
Db 1381 PLMFRCSTVSLSDFERSSTASSVQSPCGSMVSGIISPSDLDPSPGOTMPPRSKTPP 1440  
QY 1441 PPPQTAQTKREVPKNAKPTAEKRESGPKQAAVNAVQVQLPDADTLHLHFAESTPDGF 1500  
Db 1441 PPPQTAQTKREVPKNAKPTAEKRESGPKQAAVNAVQVQLPDADTLHLHFAESTPDGF 1500  
QY 1501 SCSSLSLSALDEPFIQKDELRLIMPVQENDNGNETESEOPKESNENQKEAEKTIIDSE 1560  
Db 1501 SCSSLSLSALDEPFIQKDELRLIMPVQENDNGNETESEOPKESNENQKEAEKTIIDSE 1560  
QY 1561 KDLDDSDDDDIIEILEECIIISAMPTKSSRKAKPAQTASKLPPPVARKPSQLPVYKLLPS 1620  
Db 1561 KDLDDSDDDDIIEILEECIIISAMPTKSSRKAKPAQTASKLPPPVARKPSQLPVYKLLPS 1620  
QY 1621 QNRLOPQHVSFTPDGDDMPRVYCVEGTPIINFSTATSLSDLTIESPPNELAAGEVGRGAQ 1680  
Db 1621 QNRLOPQHVSFTPDGDDMPRVYCVEGTPIINFSTATSLSDLTIESPPNELAAGEVGRGAQ 1680  
QY 1681 SGEFEKRTDPIPEGRTSTDEAOGCKTSSVTIPELDNKAEBGDIILAEACINSAMPKGSHP 1740  
Db 1681 SGEFEKRTDPIPEGRTSTDEAOGCKTSSVTIPELDNKAEBGDIILAEACINSAMPKGSHP 1740  
QY 1741 FRVKIMDQVQOASASSAPNKNQLDGKKKPTSPVKPIPQNTYRTRVRKNADSKNNLN 1800  
Db 1741 FRVKIMDQVQOASASSAPNKNQLDGKKKPTSPVKPIPQNTYRTRVRKNADSKNNLN 1800  
QY 1801 AERFSDNKSQKONLKNNSKDFNDKLPNNEDVRGSPAFDPSPHYTPIEGTPIYCFSRND 1860  
Db 1801 AERFSDNKSQKONLKNNSKDFNDKLPNNEDVRGSPAFDPSPHYTPIEGTPIYCFSRND 1860  
QY 1861 SLSSLDFFDDDDVLDLREKAELKAKENKESAKVTSHTELTSNQOSANKTOAIKQPINR 1920  
Db 1861 SLSSLDFFDDDDVLDLREKAELKAKENKESAKVTSHTELTSNQOSANKTOAIKQPINR 1920  
QY 1921 GQPKPILQKSTFPOSSKDIIPRGAATDEKLNQFAIENTPVCFSNSSLSDIDQENN 1980  
Db 1921 GQPKPILQKSTFPOSSKDIIPRGAATDEKLNQFAIENTPVCFSNSSLSDIDQENN 1980  
QY 1981 NXENEPKETEPPDQGBPSKFOASGYAPKSPHVEDTVPVCFSRNSSLSDIDSDDLQ 2040  
Db 1981 NXENEPKETEPPDQGBPSKFOASGYAPKSPHVEDTVPVCFSRNSSLSDIDSDDLQ 2040  
QY 2041 ECISAMPKPKKPSRLKGNEXHSPRNMGIIIGEDLTLDLKDIOQPDSEHGLSPDSENF 2100  
Db 2041 ECISAMPKPKKPSRLKGNEXHSPRNMGIIIGEDLTLDLKDIOQPDSEHGLSPDSENF 2100  
QY 2101 WKAIQEGANSIVSSLUHQAARAAACLRQASDSDSLSKSGISLSPHLPDQEEKPT 2160  
Db 2101 WKAIQEGANSIVSSLUHQAARAAACLRQASDSDSLSKSGISLSPHLPDQEEKPT 2160  
QY 2161 SNKGPRILKPGKSTLTETKIESESGIKGGKVKYKSLITGVRNSENSEISGQMKQLOAN 2220  
Db 2161 SNKGPRILKPGKSTLTETKIESESGIKGGKVKYKSLITGVRNSENSEISGQMKQLOAN 2220  
QY 2221 MFSISGRWTMIHPGVRNSSSTSPVSKGPKLPKTPASKSPSEGOTATTSRGAKPSVKS 2280  
Db 2221 MFSISGRWTMIHPGVRNSSSTSPVSKGPKLPKTPASKSPSEGOTATTSRGAKPSVKS 2280  
QY 2281 ELSPVARQTSQIGGSSKAPSRGSRDSTPSRPAQOPLSRPIOSPGRNISTPGRNGISPPN 2340  
Db 2281 ELSPVARQTSQIGGSSKAPSRGSRDSTPSRPAQOPLSRPIOSPGRNISTPGRNGISPPN 2340

QY 2341 KLSQLPRTSPSTASTKSSGSGKMSYTSFGROMSQOQLTKOTGLSKNASSIPRSESASKG 2400  
Db 2341 KLSQLPRTSPSTASTKSSGSGKMSYTSFGROMSQOQLTKOTGLSKNASSIPRSESASKG 2400  
QY 2401 LNOVXNGANGANKVELSRMSTKSSGSESDRSERPVLVRQSTFIKEAPSTLRRKLEESA 2460  
Db 2401 LNOVXNGANGANKVELSRMSTKSSGSESDRSERPVLVRQSTFIKEAPSTLRRKLEESA 2460  
QY 2461 SFESLSPPSRPASPTRSQAQTPVLSPLPDMSLSTHSSVQAGGWRKLPNLSPTIEYNDG 2520  
Db 2461 SFESLSPPSRPASPTRSQAQTPVLSPLPDMSLSTHSSVQAGGWRKLPNLSPTIEYNDG 2520  
QY 2521 RPAKRHDIAKSHSESPSRIPINRSGTWKREHSHSSSLPRVSTWRTTSGSSSILSASSES 2580  
Db 2521 RPAKRHDIAKSHSESPSRIPINRSGTWKREHSHSSSLPRVSTWRTTSGSSSILSASSES 2580  
QY 2581 SEKAKSEDEKHVNSISGTQSKENOVSAKGTWRKIKENEFSPNTNSTQTVSSGATNGAES 2640  
Db 2581 SEKAKSEDEKHVNSISGTQSKENOVSAKGTWRKIKENEFSPNTNSTQTVSSGATNGAES 2640  
QY 2641 KTLIYQWAPAVSKTEDVWVRIEDCPINNPRSGRSPGTGNTPPVJDSVSEKANPNIKOSKDN 2700  
Db 2641 KTLIYQWAPAVSKTEDVWVRIEDCPINNPRSGRSPGTGNTPPVJDSVSEKANPNIKOSKDN 2700  
QY 2701 QAKONTVNGSVPMRTVGLNRLNSFIQVADAPDOKGTIIPGQNNPVPVSTNESSIVERT 2760  
Db 2701 QAKONTVNGSVPMRTVGLNRLNSFIQVADAPDOKGTIIPGQNNPVPVSTNESSIVERT 2760  
QY 2761 PFSSSSSSKHSSPSGTVAAARVTPFNYPNPSRKSADSTARPQOIPTPVNNNTKGRDST 2820  
Db 2761 PFSSSSSSKHSSPSGTVAAARVTPFNYPNPSRKSADSTARPQOIPTPVNNNTKGRDST 2820  
QY 2821 DSTESSGTQSPKHSGLYVTSV 2843  
Db 2821 DSTESSGTQSPKHSGLYVTSV 2843

Search completed: August 25, 2004, 17:15:10  
Job time : 195.5 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 25, 2004, 17:23:12 ; Search time 190.5 Seconds  
(without alignments)  
4695.247 Million cell updates/sec

Title: US-09-442-489F-7

Perfect score: 14566

Sequence: 1 MAASVDQLLKQVEALKMEN.....ESSGTQSPKRGSGYLTVS 2843

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1297172 seqs, 314612898 residues

Total number of hits satisfying chosen parameters: 1297172

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Published Applications AA:\*
- 1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep:\*
  - 2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep:\*
  - 3: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep:\*
  - 4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep:\*
  - 5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep:\*
  - 6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep:\*
  - 7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep:\*
  - 8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep:\*
  - 9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep:\*
  - 10: /cgn2\_6/ptodata/2/pubpaa/US09E\_PUBCOMB.pep:\*
  - 11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep:\*
  - 12: /cgn2\_6/ptodata/2/pubpaa/US09 NEW\_PUB.pep:\*
  - 13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep:\*
  - 14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep:\*
  - 15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep:\*
  - 16: /cgn2\_6/ptodata/2/pubpaa/US10 NEW\_PUB.pep:\*
  - 17: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep:\*
  - 18: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	14566	100.0	2843	9	US-09-987-482-1
2	14566	100.0	2843	12	US-10-392-113-21
3	14557	99.9	2843	16	US-10-408-765A-1970
4	14539	99.8	2843	8	US-08-681-219-32
5	14539	99.8	2843	11	US-09-230-111C-30
6	14539	99.8	2843	14	US-10-092-138-30
7	14484.5	99.4	2844	12	US-10-267-502-370
8	13178	90.5	2845	12	US-10-267-502-372
9	4738	32.5	912	9	US-09-987-482-2
10	3892	26.7	767	9	US-09-987-482-3
11	3512.5	24.1	2274	12	US-10-267-502-373
12	3415.5	23.4	2303	12	US-10-267-502-371
13	1765	12.1	2417	12	US-10-267-502-368
14	1652	11.3	325	12	US-09-915-307-5
15	1477	10.1	1633	14	US-10-029-386-33090

ALIGNMENTS

RESULT 1

US-09-987-482-1  
; Sequence 1, Application US/09987482  
; Publication No. US20020184656A1  
; GENERAL INFORMATION:  
; APPLICANT: BHANDARI, POONAM  
; APPLICANT: SHASHIDHARA, L.S.  
; TITLE OF INVENTION: IN VIVO ASSAY SYSTEM FOR SCREENING AND VALIDATION OF  
; TITLE OF INVENTION: DRUGS AND OTHER SUBSTANCES  
; FILE REFERENCE: 056859-0134  
; CURRENT APPLICATION NUMBER: US/09/987,482  
; CURRENT FILING DATE: 2002-03-21  
; NUMBER OF SEQ ID NOS: 3  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 1  
; LENGTH: 2843  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-987-482-1

Query Match	100.0%;	Score 14566;	DB 9;	Length 2843;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 2843;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	MAASVDQLLKQVEALKMENLNROLEDNSNLTLTKEASNNKVKLKQSGSIEDEAM	60	
Db	1	MAASVDQLLKQVEALKMENLNROLEDNSNLTLTKEASNNKVKLKQSGSIEDEAM	60	
Qy	61	ASSGQIDLLERLXELNLDSSNFFGVKLRSMKLSYSGSREGSVSSRSGECSVPVPGSFPR	120	
Db	61	ASSGQIDLLERLXELNLDSSNFFGVKLRSMKLSYSGSREGSVSSRSGECSVPVPGSFPR	120	
Qy	121	RGFVNGSRESTGYLELEKERSLLDADLKBEKEKDYTAQLNLTNRDLSLTENFSL	180	
Db	121	RGFVNGSRESTGYLELEKERSLLDADLKBEKEKDYTAQLNLTNRDLSLTENFSL	180	
Qy	181	QTMWTRQLEAYEARQIRVAMEEOLGTCCODMEKQAQRIARIQOIEKDIRLIRLOLSQAT	240	

Db 181 QDTRQLQYEAEQIRVAMEBQJGTCQDVEKQAPRIARIQOIEKDIIRIQLLOSQAT 240  
 Qy 241 EAPRSQNHETGSHDAERQEQGVGEINMATSNGQGSTTRMDHETASVJSSGSTHSA 300  
 Db 241 EAPRSQNHETGSHDAERQEQGVGEINMATSNGQGSTTRMDHETASVJSSGSTHSA 300  
 Qy 301 PRLTSHLGTKEWYVYLLSMLGTHDKDMSRLLAAVSSQDSICSMROSGLPLLIQLL 360  
 Db 301 PRLTSHLGTKEWYVYLLSMLGTHDKDMSRLLAAVSSQDSICSMROSGLPLLIQLL 360  
 Qy 361 HGNDKDSVLLGNRSGKEARASAAALHNIHSPQDDKGRREIRVHLLEQIRAYCETC 420  
 Db 361 HGNDKDSVLLGNRSGKEARASAAALHNIHSPQDDKGRREIRVHLLEQIRAYCETC 420  
 Qy 421 WEQEAEHPGMDQKNPMPAPVPHQICPAVCVLMKLSFDEEHRHANVELGGQATAEILQ 480  
 Db 421 WEQEAEHPGMDQKNPMPAPVPHQICPAVCVLMKLSFDEEHRHANVELGGQATAEILQ 480  
 Qy 481 VDCMYGLTNDHYSITLRRVAGMALTNLTFGDVANKATILCSMKGCMRALVAQLKSEEDL 540  
 Db 481 VDCMYGLTNDHYSITLRRVAGMALTNLTFGDVANKATILCSMKGCMRALVAQLKSEEDL 540  
 Qy 541 QOVIASVLRNLSWRADVNSKTLREVGSVKALMECALEVKESTLKSVALWNLSAHCT 600  
 Db 541 QOVIASVLRNLSWRADVNSKTLREVGSVKALMECALEVKESTLKSVALWNLSAHCT 600  
 Qy 601 ENKADICAVGALAFUVTLYTSQNTLAIIESGGILRNVSLLIATNEDHQILRENN 660  
 Db 601 ENKADICAVGALAFUVTLYTSQNTLAIIESGGILRNVSLLIATNEDHQILRENN 660  
 Qy 661 CLQTLQLKSHSLTIVSNACGLTWNLSARNPKDQBALMDMGAVSMKNLIHSHKHQIAM 720  
 Db 661 CLQTLQLKSHSLTIVSNACGLTWNLSARNPKDQBALMDMGAVSMKNLIHSHKHQIAM 720  
 Qy 721 GSAALANLWANRPKAKYDANIMSPGSSLPSEHVRKOKALEADLAHSETFDNIDNLS 780  
 Db 721 GSAALANLWANRPKAKYDANIMSPGSSLPSEHVRKOKALEADLAHSETFDNIDNLS 780  
 Qy 781 PKASHRSKQKHQKSLYGDYVDFNRHDDNRSDNFNTGNMTVLSPYLTNTVLPSSSSSRGS 840  
 Db 781 PKASHRSKQKHQKSLYGDYVDFNRHDDNRSDNFNTGNMTVLSPYLTNTVLPSSSSSRGS 840  
 Qy 841 LDSSRSKDRSLERERIGLGNVHPATENPGTSSKRGLOISTTAAQIAKWEVSAIHTS 900  
 Db 841 LDSSRSKDRSLERERIGLGNVHPATENPGTSSKRGLOISTTAAQIAKWEVSAIHTS 900  
 Qy 901 QEDRSSGTTLHCVTDERNALRRSSAAHSTNTYNTFTKSENNRCTCMPYAKLEYKRS 960  
 Db 901 QEDRSSGTTLHCVTDERNALRRSSAAHSTNTYNTFTKSENNRCTCMPYAKLEYKRS 960  
 Qy 961 NDSLNVSSSDGYKRGQMKPSIESYSEDDKSPCSYQVYPADLAHIIHANMDNDGE 1020  
 Db 961 NDSLNVSSSDGYKRGQMKPSIESYSEDDKSPCSYQVYPADLAHIIHANMDNDGE 1020  
 Qy 1021 LDPTINYSKYDQNLNRSQSPQNERWARPKHIIEDKIQEQSQNRQSTTYPVYTE 1080  
 Db 1021 LDPTINYSKYDQNLNRSQSPQNERWARPKHIIEDKIQEQSQNRQSTTYPVYTE 1080  
 Qy 1081 STDDKHLKFQFHQOQECVSPYRSRANGSETNRVGSNHGINQVQSQCQEDDYEDDKP 1140  
 Db 1081 STDDKHLKFQFHQOQECVSPYRSRANGSETNRVGSNHGINQVQSQCQEDDYEDDKP 1140  
 Qy 1141 TNYSERYSEBQHEEERPTNYSIKYNEEKHVDQPIDYSLKYATDIPSSQKQSFPSKS 1200  
 Db 1141 TNYSERYSEBQHEEERPTNYSIKYNEEKHVDQPIDYSLKYATDIPSSQKQSFPSKS 1200  
 Qy 1201 SSGQSKTEHMSSSSENTSTPSSNAKQNLHPSSAQSRSGQPKAATCKVSSINQETIQ 1260  
 Db 1201 SSGQSKTEHMSSSSENTSTPSSNAKQNLHPSSAQSRSGQPKAATCKVSSINQETIQ 1260  
 Qy 1261 TYCVEDTPICFSRCSLSLSAABDEIGCQNTQEADSAANTLOIAIEKEKIGTSAEDPV 1320  
 Db 1261 TYCVEDTPICFSRCSLSLSAABDEIGCQNTQEADSAANTLOIAIEKEKIGTSAEDPV 1320

Qy 1321 SEVPAVSHQPRTKSRRLQSSLSSESARHKAVERPSSCAKSPSKSGAOTPKSPPEHYVOET 1380  
 Db 1321 SEVPAVSHQPRTKSRRLQSSLSSESARHKAVERPSSCAKSPSKSGAOTPKSPPEHYVOET 1380  
 Qy 1381 PLMFSRCTSVSSLDSPFESRSIASSVQSEPCSGMVSGIISPSDLPPSPQTMPPSRSKTTP 1440  
 Db 1381 PLMFSRCTSVSSLDSPFESRSIASSVQSEPCSGMVSGIISPSDLPPSPQTMPPSRSKTTP 1440  
 Qy 1441 PPQOTAKTRVPPKNAKPAETAKRESGPKQAAVNAVQVLPDADTLLHFAFESTPDGF 1500  
 Db 1441 PPQOTAKTRVPPKNAKPAETAKRESGPKQAAVNAVQVLPDADTLLHFAFESTPDGF 1500  
 Qy 1501 SCSSLSALSDEFFIOKDVLRIMPPVQENDNGNETESQPKESNENQEKAEKTTDSE 1560  
 Db 1501 SCSSLSALSDEFFIOKDVLRIMPPVQENDNGNETESQPKESNENQEKAEKTTDSE 1560  
 Qy 1561 KOLLDDDDDDIEILEBIIISAMPTKSSRAKAKPAQATASKLPPPVARKPSQLPVYKLLPS 1620  
 Db 1561 KOLLDDDDDDIEILEBIIISAMPTKSSRAKAKPAQATASKLPPPVARKPSQLPVYKLLPS 1620  
 Qy 1621 QNRLOPQKHVSFTPGDDMPRVYCVEGTPIINFSTATSLDITIESPPNELAAGEGVRGAQ 1680  
 Db 1621 QNRLOPQKHVSFTPGDDMPRVYCVEGTPIINFSTATSLDITIESPPNELAAGEGVRGAQ 1680  
 Qy 1681 SGEFEKRDITPTEGRSTDEAOGGKTSSVTIPELDNKAEGDILAEICINSAMPKGSHPK 1740  
 Db 1681 SGEFEKRDITPTEGRSTDEAOGGKTSSVTIPELDNKAEGDILAEICINSAMPKGSHPK 1740  
 Qy 1741 FRVKIINDVOQOASASSAPNKNQOLDGKKKPTSPVKPIQNTETRTVRKNADSKNNLN 1800  
 Db 1741 FRVKIINDVOQOASASSAPNKNQOLDGKKKPTSPVKPIQNTETRTVRKNADSKNNLN 1800  
 Qy 1801 AERVFSDNKSQKQNLKNNSKDFNDKLPNNEDRVGSAFDPSPHYTPIEGTPICFSRND 1860  
 Db 1801 AERVFSDNKSQKQNLKNNSKDFNDKLPNNEDRVGSAFDPSPHYTPIEGTPICFSRND 1860  
 Qy 1861 SLSLDFDDDDVDLSREKAEIRKAKENKSEPAKVTSHTELTSNQQSANKTQATAKQPINR 1920  
 Db 1861 SLSLDFDDDDVDLSREKAEIRKAKENKSEPAKVTSHTELTSNQQSANKTQATAKQPINR 1920  
 Qy 1921 GQPKILOKQSTTFQSSKDIIPDRGAATDEKLQNFALIENTPVCFSHNSLSLSLSDIDQENN 1980  
 Db 1921 GQPKILOKQSTTFQSSKDIIPDRGAATDEKLQNFALIENTPVCFSHNSLSLSLSDIDQENN 1980  
 Qy 1981 NKENEPIKETEPDPSQSGEPSPQASGVAPKSFHVEDTPVCFSRNSLSLSLSDIDEDLLQ 2040  
 Db 1981 NKENEPIKETEPDPSQSGEPSPQASGVAPKSFHVEDTPVCFSRNSLSLSLSDIDEDLLQ 2040  
 Qy 2041 ECISAMPEKKKPSRLKGDNEKHSFRNMGGLGEBDLTLDLKDIOKPDSEHGLSPDSNFND 2100  
 Db 2041 ECISAMPEKKKPSRLKGDNEKHSFRNMGGLGEBDLTLDLKDIOKPDSEHGLSPDSNFND 2100  
 Qy 2101 WKAIQEGANSIVSSLHQAAAAAALSRQASDSDSLKSGISILSGSPFHLLTPDQEEKPFT 2160  
 Db 2101 WKAIQEGANSIVSSLHQAAAAAALSRQASDSDSLKSGISILSGSPFHLLTPDQEEKPFT 2160  
 Qy 2161 SNKGPRILLKPEKSTLETFTKTESKGIKGGKVKYKSLITGKVSNSSEI SQMKQPIQAN 2220  
 Db 2161 SNKGPRILLKPEKSTLETFTKTESKGIKGGKVKYKSLITGKVSNSSEI SQMKQPIQAN 2220  
 Qy 2221 MPISRGRTMIHIFGVNRNSSSTSPVSKGKPLKTPASKSPSEGTATTSPRGAKPSVKS 2280  
 Db 2221 MPISRGRTMIHIFGVNRNSSSTSPVSKGKPLKTPASKSPSEGTATTSPRGAKPSVKS 2280  
 Qy 2281 ELSPVARQTSQIGSSSKAPSRSGRSDTSPSPAQOPLSRPIQSPGRNISI SPGRNGISPPN 2340  
 Db 2281 ELSPVARQTSQIGSSSKAPSRSGRSDTSPSPAQOPLSRPIQSPGRNISI SPGRNGISPPN 2340  
 Qy 2341 KLSQLPRTSSSTASTKSSGSGKMSYTSFGQMSQNLTKQTGLSKNASSIPRSESASKG 2400  
 Db 2341 KLSQLPRTSSSTASTKSSGSGKMSYTSFGQMSQNLTKQTGLSKNASSIPRSESASKG 2400

QY 2401 LNMNNGANKKVELSRMSSTKSGESDRERPVLRQSTFIKEAPSPTLRRKLEBSA 2460  
DB 2401 LNMNNGANKKVELSRMSSTKSGESDRERPVLRQSTFIKEAPSPTLRRKLEBSA 2460  
QY 2461 SPESLSPSRPASPTRSQAQTVLSPSLPDMSLSTHSSVQAGWRKLPNLSPTIYNDG 2520  
DB 2461 SPESLSPSRPASPTRSQAQTVLSPSLPDMSLSTHSSVQAGWRKLPNLSPTIYNDG 2520  
QY 2521 RPAKHDIARSHESPSRLPINRSGTWRKREHSHSSILPRVSTWRTTSGSSSILSASSES 2580  
DB 2521 RPAKHDIARSHESPSRLPINRSGTWRKREHSHSSILPRVSTWRTTSGSSSILSASSES 2580  
QY 2581 SEKAKSEDEKHVNSTSGTKQSENQVSAKGTWRKIKENEFPSTNSTSTQTVSSGATNGAES 2640  
DB 2581 SEKAKSEDEKHVNSTSGTKQSENQVSAKGTWRKIKENEFPSTNSTSTQTVSSGATNGAES 2640  
QY 2641 KTLIIQMAFAVSKTEDVWVRIEDCPINNPRSGRSGPTGNTPPVIDSVSEKANPNIKSDXDN 2700  
DB 2641 KTLIIQMAFAVSKTEDVWVRIEDCPINNPRSGRSGPTGNTPPVIDSVSEKANPNIKSDXDN 2700  
QY 2701 QAKQNVGSGVPMRTVGLNRLNSIOVDAPQKGTETKPGQNNPVPVSETNESSIVERT 2760  
DB 2701 QAKQNVGSGVPMRTVGLNRLNSIOVDAPQKGTETKPGQNNPVPVSETNESSIVERT 2760  
QY 2761 PFSSSSSKHSHSPSGTVAARVTFPNPNPSPRKSSADSTSARPSQIPTPVNNTKKRDSKT 2820  
DB 2761 PFSSSSSKHSHSPSGTVAARVTFPNPNPSPRKSSADSTSARPSQIPTPVNNTKKRDSKT 2820  
QY 2821 DTESTSGTQSPKRGSGVYLVTSV 2843  
DB 2821 DTESTSGTQSPKRGSGVYLVTSV 2843

## RESULT 2

US-10-392-113-21  
; Sequence 21, Application US/10392113  
; Publication NO. US20030224993A1  
; GENERAL INFORMATION:  
; APPLICANT: Deleu, Laurent  
; APPLICANT: Deleu, Laurent  
; TITLE OF INVENTION: COMPOSITIONS THAT INHIBIT PROLIFERATION  
; FILE REFERENCE: 21108.000503  
; CURRENT APPLICATION NUMBER: US/10/392,113  
; CURRENT FILING DATE: 2003-03-17  
; PRIOR APPLICATION NUMBER: 60/365,078  
; PRIOR FILING DATE: 2002-03-15  
; PRIOR APPLICATION NUMBER: PCI/US01/32127  
; PRIOR FILING DATE: 2001-10-15  
; PRIOR APPLICATION NUMBER: 60/239,705  
; PRIOR FILING DATE: 2000-10-12  
; NUMBER OF SEQ ID NOS: 45  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 21  
; LENGTH: 2843  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence:/Note =  
; OTHER INFORMATION: Synthetic Construct  
US-10-392-113-21

Query Match 100.0%; Score 14566; DB 12; Length 2843;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2843; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAASVDQLLKQVEALKVENSRLRQELDNNSHLTKLETEASNKEVLKQJGSTEDEAM 60  
DB 1 MAASVDQLLKQVEALKVENSRLRQELDNNSHLTKLETEASNKEVLKQJGSTEDEAM 60  
QY 61 ASSGQIDLLERKELNLDSSNFPVGLRSMKSLRSGYSGREGSVSSRSGECPVPMGSPFR 120  
DB 61 ASSGQIDLLERKELNLDSSNFPVGLRSMKSLRSGYSGREGSVSSRSGECPVPMGSPFR 120

QY 121 RGFVNGSRESTGYLLELEKERSLILLADLKEEKEKWYQAQLQNLTKRIDSLPLTFENFSL 180  
DB 121 RGFVNGSRESTGYLLELEKERSLILLADLKEEKEKWYQAQLQNLTKRIDSLPLTFENFSL 180  
QY 181 QDWTRELQLEYEARQIRVAMEBEOIGTCQDMKEKAORRIARIQOIEKDIIRIROLQSOAT 240  
DB 181 QDWTRELQLEYEARQIRVAMEBEOIGTCQDMKEKAORRIARIQOIEKDIIRIROLQSOAT 240  
QY 241 EAERSSQNKHETGSHDAERQNEGQVGEINWATSGNGCGSTTRMDHETASVLSSSSTHSA 300  
DB 241 EAERSSQNKHETGSHDAERQNEGQVGEINWATSGNGCGSTTRMDHETASVLSSSSTHSA 300  
QY 301 PRLTSLHGTQKVMVYLLSMLGTHDKDDMSRRTLLAMSSQDSCISNMRQSGCPLLIQLL 360  
DB 301 PRLTSLHGTQKVMVYLLSMLGTHDKDDMSRRTLLAMSSQDSCISNMRQSGCPLLIQLL 360  
QY 361 HGNDKDSVLLNGRSGSEARARASAAALHNIHSDPDKRGRREIRVHLLEQIRAYCETC 420  
DB 361 HGNDKDSVLLNGRSGSEARARASAAALHNIHSDPDKRGRREIRVHLLEQIRAYCETC 420  
QY 421 WEMQEAHEPGMDQKPNMPAPVEHQICPAVCVLMKLSFDEEHRHAMNELGGLQAIABLLQ 480  
DB 421 WEMQEAHEPGMDQKPNMPAPVEHQICPAVCVLMKLSFDEEHRHAMNELGGLQAIABLLQ 480  
QY 481 VDCMYGLTNDHYSITILRRYAGMALTNLTGPDYVANKATILCSMKGCMRALVAOLKSESED 540  
DB 481 VDCMYGLTNDHYSITILRRYAGMALTNLTGPDYVANKATILCSMKGCMRALVAOLKSESED 540  
QY 541 QQVIASVLRNLSWRADVNSKTLREVGSKVAKAMECALEVKEKSTLKSVALNLSAHCT 600  
DB 541 QQVIASVLRNLSWRADVNSKTLREVGSKVAKAMECALEVKEKSTLKSVALNLSAHCT 600  
QY 601 ENKADICAVDGAALAFVLGTLTYRSQNTNLAIIESGGGILRNVSLSIATNEDHQILRENN 660  
DB 601 ENKADICAVDGAALAFVLGTLTYRSQNTNLAIIESGGGILRNVSLSIATNEDHQILRENN 660  
QY 661 CLOTLLOHLKSHSLTIVSNACGTLNLSARNPKDQEAALWDMGAVSMLKNIHSHKXMIAM 720  
DB 661 CLOTLLOHLKSHSLTIVSNACGTLNLSARNPKDQEAALWDMGAVSMLKNIHSHKXMIAM 720  
QY 721 GSAAALRNLMANRPAPKYKDNIMSPGSSLPSSLHVRKQKALEAEALDAOHLSETFDNIDNLS 780  
DB 721 GSAAALRNLMANRPAPKYKDNIMSPGSSLPSSLHVRKQKALEAEALDAOHLSETFDNIDNLS 780  
QY 781 PKASHRSKQKHOSLYGQVYFDNRHDDNDSDFNTGNMTVLSPYLNTTVLPSSSSSRGS 840  
DB 781 PKASHRSKQKHOSLYGQVYFDNRHDDNDSDFNTGNMTVLSPYLNTTVLPSSSSSRGS 840  
QY 841 LSSRSSEKDRSLERERIGLGNYPATENPGTSSKRGLOISTTAAQIAKYMEEVSAIHTS 900  
DB 841 LSSRSSEKDRSLERERIGLGNYPATENPGTSSKRGLOISTTAAQIAKYMEEVSAIHTS 900  
QY 901 QEDRSSGSTTELHCVTDERNALRRSSAAHSTHTYFTKSENSENRTCSMPYAKLEYKRSS 960  
DB 901 QEDRSSGSTTELHCVTDERNALRRSSAAHSTHTYFTKSENSENRTCSMPYAKLEYKRSS 960  
QY 961 NDSLNSVSSSDGYGKRGOMKPSIESYSEDDSEKFCVGYQPADLAHKIHSANHMDNDGE 1020  
DB 961 NDSLNSVSSSDGYGKRGOMKPSIESYSEDDSEKFCVGYQPADLAHKIHSANHMDNDGE 1020  
QY 1021 LDPTINYSKYSDQNLNSGRQSPSONERWAPKHIIIEDEIKQSEQRQSRNQSTTYPVYTE 1080  
DB 1021 LDPTINYSKYSDQNLNSGRQSPSONERWAPKHIIIEDEIKQSEQRQSRNQSTTYPVYTE 1080  
QY 1081 STDDKHLKFPQHFQOQECVSPYSRGANGSETNRVGSNGHGINONVSQSLCOEDDYEDDKP 1140  
DB 1081 STDDKHLKFPQHFQOQECVSPYSRGANGSETNRVGSNGHGINONVSQSLCOEDDYEDDKP 1140  
QY 1141 TNYSERYSEQHEEERPTNYSIKYNEEKRRVDQPIDYSLKYATDIPSSQKQSFSPSKS 1200  
DB 1141 TNYSERYSEQHEEERPTNYSIKYNEEKRRVDQPIDYSLKYATDIPSSQKQSFSPSKS 1200

1201 SSGSSKTEHMSSESSENTSTPSSNAKQNLHPSSAQSRSGQPKAATCKVSSINOETIQ 1260  
1201 SSGSSKTEHMSSESSENTSTPSSNAKQNLHPSSAQSRSGQPKAATCKVSSINOETIQ 1260  
1261 TYCVEDTPICFSRCSLSLSLSAEDDEICGNOTTQEADSAANTLQIAIEIKKIGTRSAEDPV 1320  
1261 TYCVEDTPICFSRCSLSLSLSAEDDEICGNOTTQEADSAANTLQIAIEIKKIGTRSAEDPV 1320  
1321 SEVPAVQCHRTKSRRLQSSLSSESARHKAHAFSSGAKSPKSGAQTPKSPPEHYVQET 1380  
1321 SEVPAVQCHRTKSRRLQSSLSSESARHKAHAFSSGAKSPKSGAQTPKSPPEHYVQET 1380  
1381 PLMFSTRCTVSSLDSPESRISAGSVQPCSGMVSGIISPSDLPDGPOTMPPSRSKTTP 1440  
1381 PLMFSTRCTVSSLDSPESRISAGSVQPCSGMVSGIISPSDLPDGPOTMPPSRSKTTP 1440  
1441 PPQTAQTKREVPKNKAAPTAKESGPKQAAVNAAVORVQVLPDADTLHLHFAESTPDGF 1500  
1441 PPQTAQTKREVPKNKAAPTAKESGPKQAAVNAAVORVQVLPDADTLHLHFAESTPDGF 1500  
1501 SCSSLSALSLDEPFIQDVELRIMPVQENDNGNETESQPKESNENOEKAEKIDSE 1560  
1501 SCSSLSALSLDEPFIQDVELRIMPVQENDNGNETESQPKESNENOEKAEKIDSE 1560  
1561 KDLDDDDDDDDIELEBECIISAMPTKSSRAKKAQAQASKLPPPVARKPSQLPVYKLLPS 1620  
1561 KDLDDDDDDDDIELEBECIISAMPTKSSRAKKAQAQASKLPPPVARKPSQLPVYKLLPS 1620  
1621 QNRLOPKHVSFTPGDDMPVYCVGFTPIINFSTATSLDITIESPNELAAEGVGGQA 1680  
1621 QNRLOPKHVSFTPGDDMPVYCVGFTPIINFSTATSLDITIESPNELAAEGVGGQA 1680  
1681 SGEFEKDDTPTTEGRSTDEAGGKTSSVTIPELDDNKAEGDILAEICINSAMPKGSHKP 1740  
1681 SGEFEKDDTPTTEGRSTDEAGGKTSSVTIPELDDNKAEGDILAEICINSAMPKGSHKP 1740  
1741 FRVKKIMDVQQAASASAPNKNQLDGKKKPTSPVKPIQNTTEYTRVRKNADSKNLN 1800  
1741 FRVKKIMDVQQAASASAPNKNQLDGKKKPTSPVKPIQNTTEYTRVRKNADSKNLN 1800  
1801 AERFSDNKSQKQNLKNSKDFNDKLPNNEDVRGSAFDSPHHTPTTEGTPVCFSRND 1860  
1801 AERFSDNKSQKQNLKNSKDFNDKLPNNEDVRGSAFDSPHHTPTTEGTPVCFSRND 1860  
1861 SLSSLDPDDDDVLSREKAEARAKENKESAKVTSTELTSNQOQANKTQAIKQPINR 1920  
1861 SLSSLDPDDDDVLSREKAEARAKENKESAKVTSTELTSNQOQANKTQAIKQPINR 1920  
1921 GQPKPILOKQSTPQSSKQIPDRGAATDEKLQFAIENTVCFSHNSLSLSLSDIDQENN 1980  
1921 GQPKPILOKQSTPQSSKQIPDRGAATDEKLQFAIENTVCFSHNSLSLSLSDIDQENN 1980  
1981 NKNEPIKETEPDPSQOQEPKQASGAPKSFHVEDTPVCFSRNSLSLSLSDIDDLQ 2040  
1981 NKNEPIKETEPDPSQOQEPKQASGAPKSFHVEDTPVCFSRNSLSLSLSDIDDLQ 2040  
2041 ECISAMPKPKKRLKGDNEKSPRNGGILGEDTLTDKIDQRPDSEHGLSPDENFD 2100  
2041 ECISAMPKPKKRLKGDNEKSPRNGGILGEDTLTDKIDQRPDSEHGLSPDENFD 2100  
2101 WKALQEGANSIVSLHQAACLSRAQSSDSLSLSKSGISLGSPFHLTPQEEKPKPT 2160  
2101 WKALQEGANSIVSLHQAACLSRAQSSDSLSLSKSGISLGSPFHLTPQEEKPKPT 2160  
2161 SNKGPRILKPGKSTLTETKIESKIGKGGKYKSLITGKVRNSSEISGQWKQPLQAN 2220  
2161 SNKGPRILKPGKSTLTETKIESKIGKGGKYKSLITGKVRNSSEISGQWKQPLQAN 2220  
2221 MPSTSRGRMTIHIPGVNRSSSTSPVSKGPPPLKTPASKSPSQGTATTPRGAKSQVKS 2280  
2221 MPSTSRGRMTIHIPGVNRSSSTSPVSKGPPPLKTPASKSPSQGTATTPRGAKSQVKS 2280  
2281 ELSPVARTSQIGGSSKAPSRSGRSDTSPRPAQQLSRPIQSPGRNSISPGRNGISPPN 2340

2281 ELSPVARTSQIGGSSKAPSRSGRSDTSPRPAQQLSRPIQSPGRNSISPGRNGISPPN 2340  
2341 KLSQLPRTSSPTASTKSSGSKMYTSPGRMQQNLTKQTGLSKNASSIPRSESASKG 2400  
2341 KLSQLPRTSSPTASTKSSGSKMYTSPGRMQQNLTKQTGLSKNASSIPRSESASKG 2400  
2401 LQNMNGANGANKVELSRMSTKSSGSDRSEBVPVLRQSTFIKEAPSPTLRKLSESA 2460  
2401 LQNMNGANGANKVELSRMSTKSSGSDRSEBVPVLRQSTFIKEAPSPTLRKLSESA 2460  
2461 SFESLSPSSRPASPTRSQAOQTPVLSPLPDMSLSTHSSVQAGGWRKLPNLSPTIEYNDG 2520  
2461 SFESLSPSSRPASPTRSQAOQTPVLSPLPDMSLSTHSSVQAGGWRKLPNLSPTIEYNDG 2520  
2521 RPAKEHDIARSHSPSRPLPINRSGTWKREHSHSSSLPVSTWRRTGSSSSILSASSES 2580  
2521 RPAKEHDIARSHSPSRPLPINRSGTWKREHSHSSSLPVSTWRRTGSSSSILSASSES 2580  
2581 SEKAKSEDEKHVNSISGTSKQENQVSAKGTWRKIKENEFSPNTNSTQTSVSSGATNGAES 2640  
2581 SEKAKSEDEKHVNSISGTSKQENQVSAKGTWRKIKENEFSPNTNSTQTSVSSGATNGAES 2640  
2641 KTLIQMAPAVSKTEDVWRIEDCPINNPRSGRSPGTGNTPPVIDSVSEKANPNIKDSKN 2700  
2641 KTLIQMAPAVSKTEDVWRIEDCPINNPRSGRSPGTGNTPPVIDSVSEKANPNIKDSKN 2700  
2701 QAKQVNGSGVPMRTVGLNRLNSFIOVDAPDQKGTIEIKQGNPNVPVSEINNESSIVERT 2760  
2701 QAKQVNGSGVPMRTVGLNRLNSFIOVDAPDQKGTIEIKQGNPNVPVSEINNESSIVERT 2760  
2761 PFSSSSSKHSSPSTVAARVTPNPNPSPKSADSTARSPOIPTPVNNNTKKRUSKT 2820  
2761 PFSSSSSKHSSPSTVAARVTPNPNPSPKSADSTARSPOIPTPVNNNTKKRUSKT 2820  
2821 DSTESEGTQSPKSHSGSYLTVSV 2843  
2821 DSTESEGTQSPKSHSGSYLTVSV 2843

RESULT 3  
US-10-408-765A-1970  
; Sequence 1970, Application US/10408765A  
; Publication No. US20040101874A1  
; GENERAL INFORMATION:  
; APPLICANT: Ghosh, Soumitra S.  
; APPLICANT: Fahy, Boia D.  
; APPLICANT: Zhang, Bing  
; APPLICANT: Gibson, Bradford W.  
; APPLICANT: Taylor, Steven W.  
; APPLICANT: Glenn, Gary M.  
; APPLICANT: Warnock, Dale E.  
; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION  
; TITLE OF INVENTION: IDENTIFIED IN THE MITOCHONDRIAL PROTEOME  
; FILE REFERENCE: 660088.465  
; CURRENT APPLICATION NUMBER: US/10/408,765A  
; CURRENT FILING DATE: 2003-04-04  
; NUMBER OF SEQ ID NOS: 3077  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1970  
; LENGTH: 2843  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-408-765A-1970

Query Match 99.9%; Score 14557; DB 16; Length 2843;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2842; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MAASVYDQLLKQVEALKMNSNLQCELDNNSHLTKLETSANMKVILKQLQGSIEDEAM 60  
DB 1 MAASVYDQLLKQVEALKMNSNLQCELDNNSHLTKLETSANMKVILKQLQGSIEDEAM 60



QY 61 ASSGQIDLLERKELNLDSSNPPGKLRKMSLRYSYSGREGSVSRSGECPVPMWGFPR 120  
DB 61 ASSGQIDLLERKELNLDSSNPPGKLRKMSLRYSYSGREGSVSRSGECPVPMWGFPR 120  
QY 121 RGfVNGSREStCYLBELEKERSLLADLDKEEKEDWYIAQI ONLTXRIDS:PLTENSFL 180  
DB 121 RGfVNGSREStCYLBELEKERSLLADLDKEEKEDWYIAQI ONLTXRIDS:PLTENSFL 180  
QY 181 QTDMTRRQLEYEARQIRVAMEEQLGTCQDMKRAQRRIAR:QQIEKDIRRQLLOSOAT 240  
DB 181 QTDMTRRQLEYEARQIRVAMEEQLGTCQDMKRAQRRIAR:QQIEKDIRRQLLOSOAT 240  
QY 241 EAERSSQKHETGSHDAERONEGQGVGEINMATSGOGSTTRMDHETASVLSSSSTHSA 300  
DB 241 EAERSSQKHETGSHDAERONEGQGVGEINMATSGOGSTTRMDHETASVLSSSSTHSA 300  
QY 301 PERLTSHLGTKEVWYSLMLGTHDKDDMERTLLAMSSSDSCISMOSQCLPLLIQLL 360  
DB 301 PERLTSHLGTKEVWYSLMLGTHDKDDMERTLLAMSSSDSCISMOSQCLPLLIQLL 360  
QY 361 HGNDXSVLLGNRSRGSKEARAPASAAALHNI:THSQPDDKRGREIRVZLHLBQIRAYCETC 420  
DB 361 HGNDXSVLLGNRSRGSKEARAPASAAALHNI:THSQPDDKRGREIRVZLHLBQIRAYCETC 420  
QY 421 NEWQSAHEFGMDQDNMPAPVEHQCIPAVCVLMKLSFDEBHRHAMNELGGLQIAELLO 480  
DB 421 NEWQSAHEFGMDQDNMPAPVEHQCIPAVCVLMKLSFDEBHRHAMNELGGLQIAELLO 480  
QY 481 VDCEMYGLTNDHYSITLARRYAGMALTNLTFGDVANKATLCMKGCMRALVAQLKSESDL 540  
DB 481 VDCEMYGLTNDHYSITLARRYAGMALTNLTFGDVANKATLCMKGCMRALVAQLKSESDL 540  
QY 541 QOVIASVLRNLISWRADVNSKTLREVSVKALMECALEVKESTLKSLSALWNLISAHCT 600  
DB 541 QOVIASVLRNLISWRADVNSKTLREVSVKALMECALEVKESTLKSLSALWNLISAHCT 600  
QY 601 ENKADICAVDGALEFLVGLTYRSQTNLTALIESGGGILRVNVSIIATNEDHROILRENN 660  
DB 601 ENKADICAVDGALEFLVGLTYRSQTNLTALIESGGGILRVNVSIIATNEDHROILRENN 660  
QY 661 CLQTLQHLKSHSLTIVSNACOTLWNLISARNPKDQEALWDMGAVSMLKXNL:HSKHKMIAM 720  
DB 661 CLQTLQHLKSHSLTIVSNACOTLWNLISARNPKDQEALWDMGAVSMLKXNL:HSKHKMIAM 720  
QY 721 GSAALRLNMANRPKADANIMSPGSSILPSLHVTKOKALEAELDAOHLSETFNIDNLS 780  
DB 721 GSAALRLNMANRPKADANIMSPGSSILPSLHVTKOKALEAELDAOHLSETFNIDNLS 780  
QY 781 PRASHRSQRHQRKQSLYGDYVFDTRNHRDNRSDNFNTGNMTVLSPLYINTTVLPSSSSSRGS 840  
DB 781 PRASHRSQRHQRKQSLYGDYVFDTRNHRDNRSDNFNTGNMTVLSPLYINTTVLPSSSSSRGS 840  
QY 841 LDRSSEKDRSLERERGLGNHYHATENPGTSSKRGLOISTTAAQIAKWEEVSAIHTS 900  
DB 841 LDRSSEKDRSLERERGLGNHYHATENPGTSSKRGLOISTTAAQIAKWEEVSAIHTS 900  
QY 901 QDRSGSGTTELHCVTDENALRRSSAAHTHNTYNTFKSENSNRNFTCMYPKALEYKRSS 960  
DB 901 QDRSGSGTTELHCVTDENALRRSSAAHTHNTYNTFKSENSNRNFTCMYPKALEYKRSS 960  
QY 961 NDSLNSVSSSDGYKRGQWKPISYSDSDDESKFCSYGOYPADLAHK:HSANHMDNDNDE 1020  
DB 961 NDSLNSVSSSDGYKRGQWKPISYSDSDDESKFCSYGOYPADLAHK:HSANHMDNDNDE 1020  
QY 1021 LPTPINYSUKYDEQNLNSGRQSPQNERWARPKHIEDEIKOSEQRQSRNOSTTYPVYTE 1080  
DB 1021 LPTPINYSUKYDEQNLNSGRQSPQNERWARPKHIEDEIKOSEQRQSRNOSTTYPVYTE 1080  
QY 1081 STDDKHLKFPQHFQGOECVPSYRSGANGETNRVGSNHGINQVNSQSLCQEDDDYEDDKP 1140  
DB 1081 STDDKHLKFPQHFQGOECVPSYRSGANGETNRVGSNHGINQVNSQSLCQEDDDYEDDKP 1140  
QY 1141 TNYSEYSEEEQHEERPTNYSIKYNEEKRHVDQPIDYSLKYATDIPSSQKQSFSPKS 1200

DB 1141 TNYSEYSEEEQHEERPTNYSIKYNEEKRHVDQPIDYSLKYATDIPSSQKQSFSPKS 1200  
QY 1201 SSGOSSKTEHSSSSSNTSTPSSNAKRONQLHPSSSAQSSGQPKAATCKVSSINQETIQ 1260  
DB 1201 SSGOSSKTEHSSSSSNTSTPSSNAKRONQLHPSSSAQSSGQPKAATCKVSSINQETIQ 1260  
QY 1261 TYCVETPTICFSCSSLSLSAEDIEIGNQTTQEADSANTLQIABIKIGKIGRSADPV 1320  
DB 1261 TYCVETPTICFSCSSLSLSAEDIEIGNQTTQEADSANTLQIABIKIGKIGRSADPV 1320  
QY 1321 SEVPAYSOHPRTKSSRLQSSLSSESARHKAFFSSGAKSPSKSGAQTPKSPHHVVOET 1380  
DB 1321 SEVPAYSOHPRTKSSRLQSSLSSESARHKAFFSSGAKSPSKSGAQTPKSPHHVVOET 1380  
QY 1381 PLMFSCRTSVSSLDSPESRSIASSVOSEPCSGMVSGIISPSDLDPSPGQTMPPSRSKTTP 1440  
DB 1381 PLMFSCRTSVSSLDSPESRSIASSVOSEPCSGMVSGIISPSDLDPSPGQTMPPSRSKTTP 1440  
QY 1441 PPTQATQKREVPKNKAPTAEKRESGPKQAANVAQVQVLPDADTLHLHFAESTPDGF 1500  
DB 1441 PPTQATQKREVPKNKAPTAEKRESGPKQAANVAQVQVLPDADTLHLHFAESTPDGF 1500  
QY 1501 SCSSLSALSALDPEFFIQKOVLEIRIMPPVQENDNGNTESEQPKESNEQKEAEKIDSE 1560  
DB 1501 SCSSLSALSALDPEFFIQKOVLEIRIMPPVQENDNGNTESEQPKESNEQKEAEKIDSE 1560  
QY 1561 KDLLDSDDDDDIILBECIISAMPTKSSRKAKPAQATASKLPPPVARKPSQLPVYKLLPS 1620  
DB 1561 KDLLDSDDDDDIILBECIISAMPTKSSRKAKPAQATASKLPPPVARKPSQLPVYKLLPS 1620  
QY 1621 QNRLOPKHVSFTPGDDMPREVYCVETGTPINFSTATSLDLTIESPPNELAAGVGVRGAQ 1680  
DB 1621 QNRLOPKHVSFTPGDDMPREVYCVETGTPINFSTATSLDLTIESPPNELAAGVGVRGAQ 1680  
QY 1681 SGFEKEDTPTTEGRSTDEAQQGKTSVTIPELDDNKAEBGDILAEICINSAMPKGSKHP 1740  
DB 1681 SGFEKEDTPTTEGRSTDEAQQGKTSVTIPELDDNKAEBGDILAEICINSAMPKGSKHP 1740  
QY 1741 FRVKKIMQVQOASASSAPKNQOLGKKKTPSPVKPIPONTETRYTRVKNADSKNLN 1800  
DB 1741 FRVKKIMQVQOASASSAPKNQOLGKKKTPSPVKPIPONTETRYTRVKNADSKNLN 1800  
QY 1801 AERFSDNKDKQKQNLKNKDNKDFNPNNEDRVRSFAPDSPHVYTPTEGTPYCFSRND 1860  
DB 1801 AERFSDNKDKQKQNLKNKDNKDFNPNNEDRVRSFAPDSPHVYTPTEGTPYCFSRND 1860  
QY 1861 SLSSLDPEDDDDVLSREKABELRKAENKESAEKVTSHTELTSNQOQANKTOAIKOPINR 1920  
DB 1861 SLSSLDPEDDDDVLSREKABELRKAENKESAEKVTSHTELTSNQOQANKTOAIKOPINR 1920  
QY 1921 GQPKPILQKOSTTPOSSKDI:PDGGAATDEKLNFAIENTVPCFSHNSLSLSIDDOENN 1980  
DB 1921 GQPKPILQKOSTTPOSSKDI:PDGGAATDEKLNFAIENTVPCFSHNSLSLSIDDOENN 1980  
QY 1981 NKNEPIKETEPDPDSQGEPSKPOASGYAPKSFHVEDTPVCFSRNSLSLSIDSEDDLLQ 2040  
DB 1981 NKNEPIKETEPDPDSQGEPSKPOASGYAPKSFHVEDTPVCFSRNSLSLSIDSEDDLLQ 2040  
QY 2041 ECITSSAMPKKKPSRLKGDNEKHSPRNMGILGEDITLDKI:QRPDSEHGLSPDSNFD 2100  
DB 2041 ECITSSAMPKKKPSRLKGDNEKHSPRNMGILGEDITLDKI:QRPDSEHGLSPDSNFD 2100  
QY 2101 WKATQEGANSIVSLHQAAAAACLSRQASSDSDSILSKSGISLGSPFHLTPQEEKPFT 2160  
DB 2101 WKATQEGANSIVSLHQAAAAACLSRQASSDSDSILSKSGISLGSPFHLTPQEEKPFT 2160  
QY 2161 SNKGPRILKPGEXSTLTETKKIESKGIKGGKKVYKSLITGKVRNSEISQWKKPQLOAN 2220  
DB 2161 SNKGPRILKPGEXSTLTETKKIESKGIKGGKKVYKSLITGKVRNSEISQWKKPQLOAN 2220  
QY 2221 MPSTSRGRTMIHPIGVNNSSSSTSPVSKGPPPLTKPASKPSGQATATTSPROAKPSVKS 2280

Db 2221 MPSISRGRTMIHIPVRNSSSTSPVSKGPKLTKPASKSPSEGTATTSPRGAKPSVKS 2280  
 QY 2281 ELSVARTQIOIGSSKAPRSGSRDSTPRPAQPLSRPQSPGNSISPGNGISPN 2340  
 Db 2281 ELSVARTQIOIGSSKAPRSGSRDSTPRPAQPLSRPQSPGNSISPGNGISPN 2340  
 QY 2341 KLSQLPRTSPSTASTKSSGSGKMSYTPSGRQMSQQNLTKQTLGSKNASSIPRSESASG 2400  
 Db 2341 KLSQLPRTSPSTASTKSSGSGKMSYTPSGRQMSQQNLTKQTLGSKNASSIPRSESASG 2400  
 QY 2401 LQWNGNGANKKVELSRMSTKSSGSDRSPRVLVRQSTFIKAPSTLRKLEESA 2460  
 Db 2401 LQWNGNGANKKVELSRMSTKSSGSDRSPRVLVRQSTFIKAPSTLRKLEESA 2460  
 QY 2461 SFESLSPSSRPASPTRSQATPVLSPLPDMSLSTHSSVQAGGWRKLPNLSPTIEYNDG 2520  
 Db 2461 SFESLSPSSRPASPTRSQATPVLSPLPDMSLSTHSSVQAGGWRKLPNLSPTIEYNDG 2520  
 QY 2521 RPAKRHDIAHSHSPRLPINRSGTWKREHSHSSSLPRVSTWRTGSSSSILSASSES 2580  
 Db 2521 RPAKRHDIAHSHSPRLPINRSGTWKREHSHSSSLPRVSTWRTGSSSSILSASSES 2580  
 QY 2581 SEKAKSEDEKHVNSISGTGSKENQVSAKGTWKIKENEFSPNSTSQTVSSGATNGAES 2640  
 Db 2581 SEKAKSEDEKHVNSISGTGSKENQVSAKGTWKIKENEFSPNSTSQTVSSGATNGAES 2640  
 QY 2641 KTLIQMAPAVSKTEDVWVRIEDCPINNPRSGSPGTGNTPPVIDSVSEKANPNIKDKDN 2700  
 Db 2641 KTLIQMAPAVSKTEDVWVRIEDCPINNPRSGSPGTGNTPPVIDSVSEKANPNIKDKDN 2700  
 QY 2701 OAKQNVGSGVPMRTVGLNRLNSFTQVADPQKGTIKPQNNPVVSFTNESSIVERT 2760  
 Db 2701 OAKQNVGSGVPMRTVGLNRLNSFTQVADPQKGTIKPQNNPVVSFTNESSIVERT 2760  
 QY 2761 PFSSSSSHSSPSGTVAARVTPFNPNPSPKSSADSTSAKPSQIPTPVNNNTKKRDSKT 2820  
 Db 2761 PFSSSSSHSSPSGTVAARVTPFNPNPSPKSSADSTSAKPSQIPTPVNNNTKKRDSKT 2820  
 QY 2821 DSTESSGTQSPKHSGLYTSV 2843  
 Db 2821 DSTESSGTQSPKHSGLYTSV 2843

RESULT 4

US-08-681-219-32  
 ; Sequence 32, Application US/08681219  
 ; Publication No. US20020058607A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Takaaki Sato and Junn Yanagisawa  
 ; TITLE OF INVENTION: COMPOUNDS THAT INHIBIT THE INTERACTION BETWEEN  
 ; TITLE OF INVENTION: SIGNAL-TRANSDUCING PROTEINS AND THE GLGF  
 ; TITLE OF INVENTION: (PDZ/DHR) DOMAIN AND USES THEREOF  
 ; NUMBER OF SEQUENCES: 35  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESS: Cooper & Dunham LLP  
 ; STREET: 1185 Avenue of the Americas  
 ; CITY: New York  
 ; STATE: New York  
 ; COUNTRY: U.S.A.  
 ; ZIP: 10036  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA: US/08/681,219  
 ; FILING DATE: 22-JUL-1996  
 ; CLASSIFICATION: 435  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: White, John P  
 ; REGISTRATION NUMBER: 28,678  
 ; REFERENCE/DOCKET NUMBER: 0575/48962/JFPW/JKM

TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (212) 278-0400  
 ; TELEFAX: (212) 391-0525  
 ; INFORMATION FOR SEQ ID NO: 32:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 2843 amino acids  
 ; TYPE: amino acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: peptide  
 ; US-08-681-219-32  
 Query Match 99.8%; Score 14539; DB 8; Length 2843;  
 Best Local Similarity 99.8%; Pred. No. 0;  
 Matches 2837; Conservative 3; Mismatches 3; Indels 0; Gaps 0;  
 QY 1 MAASDYDQLLKQVEALXWENSNLKQLEEDNSNHLTKLETSANMKVYLKOLQGSIEDEAM 60  
 Db 1 MAASDYDQLLKQVEALXWENSNLKQLEEDNSNHLTKLETSANMKVYLKOLQGSIEDEAM 60  
 QY 61 ASSGQIDLLERLKEINLNDSSNFPQVKLRKMSLRSGSREGSVSSRSGECSPVPMGSPFR 120  
 Db 61 ASSGQIDLLERLKEINLNDSSNFPQVKLRKMSLRSGSREGSVSSRSGECSPVPMGSPFR 120  
 QY 121 RGFVNGSRESTGYLEELKERSLLADLDKEEKKDNYAQLQNLTKRIDSPLTENFSL 180  
 Db 121 RGFVNGSRESTGYLEELKERSLLADLDKEEKKDNYAQLQNLTKRIDSPLTENFSL 180  
 QY 181 QDTRRQLEYEARQIRVAMEEQGLTQDMKRAQRRIARIQIEKDIILRIQLQSQAT 240  
 Db 181 QDTRRQLEYEARQIRVAMEEQGLTQDMKRAQRRIARIQIEKDIILRIQLQSQAT 240  
 QY 241 BAERSSQNHKTGSHDAERQNEGGVGEINMATSGNQGSGSTTRMDHTASVLSSSSTHSA 300  
 Db 241 BAERSSQNHKTGSHDAERQNEGGVGEINMATSGNQGSGSTTRMDHTASVLSSSSTHSA 300  
 QY 301 PRRLTSHLGTQVEMVYLLSMGLTHDKDMSRTLLAMSSQDSCI SMRQSGCLPLLIQLL 360  
 Db 301 PRRLTSHLGTQVEMVYLLSMGLTHDKDMSRTLLAMSSQDSCI SMRQSGCLPLLIQLL 360  
 QY 361 HGNDKDSVLLGNRSGSKEARASAAALHNIHSGOPDDKRGRRREIRVLHLEQIRAYCETC 420  
 Db 361 HGNDKDSVLLGNRSGSKEARASAAALHNIHSGOPDDKRGRRREIRVLHLEQIRAYCETC 420  
 QY 421 WEQEAHEPGMDQKPNMPAPVEHQICPAVCVLMKLSFDEEHRHANNELGLOQAIALLQ 480  
 Db 421 WEQEAHEPGMDQKPNMPAPVEHQICPAVCVLMKLSFDEEHRHANNELGLOQAIALLQ 480  
 QY 481 VDCMEYGLTNDHYITLRYAGMALTNLTGADVANKATLCMKGCMRALVAQLKSEEDL 540  
 Db 481 VDCMEYGLTNDHYITLRYAGMALTNLTGADVANKATLCMKGCMRALVAQLKSEEDL 540  
 QY 541 QOVIASVLRNLSWADVNSKKTIREVGSVKALMECALEVKKESTLKSVLSALNLSAHT 600  
 Db 541 QOVIASVLRNLSWADVNSKKTIREVGSVKALMECALEVKKESTLKSVLSALNLSAHT 600  
 QY 601 ENKADICAVDGAFLVGTTLTYRSQTNTLAIIESGGGILRNVSLSIATNEDHRQLRENN 660  
 Db 601 ENKADICAVDGAFLVGTTLTYRSQTNTLAIIESGGGILRNVSLSIATNEDHRQLRENN 660  
 QY 661 CLQTLQHLKSHSLTIIVSNACGTLWNI SARPKDQEBALWDMGAVSMKLNLIHSHKHMAM 720  
 Db 661 CLQTLQHLKSHSLTIIVSNACGTLWNI SARPKDQEBALWDMGAVSMKLNLIHSHKHMAM 720  
 QY 721 GSAALRNLMANRPAYKDNIMSPGSSSLPSLHVKKOKALEAEADAHLSETFDNIDNLS 780  
 Db 721 GSAALRNLMANRPAYKDNIMSPGSSSLPSLHVKKOKALEAEADAHLSETFDNIDNLS 780  
 QY 781 PKASHRSKQRHKQSLYGDYVDFVTNRHDDNRSDNPNFTGNMTVLSPLYLNTTVLPSSSSSRGS 840  
 Db 781 PKASHRSKQRHKQSLYGDYVDFVTNRHDDNRSDNPNFTGNMTVLSPLYLNTTVLPSSSSSRGS 840  
 QY 841 LDSSRSKORSLEEREGIGLGNYPATENPGTSSKRGSLQISTTAAQIAKVMVEVSAHTS 900

841 LDSSRSKDRSLERERGI L G N Y H P A T E N P O T S S K R G L Q I O T T A A Q I A K W E E V S A I H T S 900  
901 QEDRSSGTTTLHCVTDERNALRSSAAHTSNYNTFKSENSNRTCSMPYAKLEYKRS 960  
901 QEDRSSGTTTLHCVTDERNALRSSAAHTSNYNTFKSENSNRTCSMPYAKLEYKRS 960  
961 NDSLNSVSSDGYGKRGQWPKPSIESYSDDESCKFCYGYQYPADLAHKIHSANHMDDNDGE 1020  
961 NDSLNSVSSDGYGKRGQWPKPSIESYSDDESCKFCYGYQYPADLAHKIHSANHMDDNDGE 1020  
1021 LDTPINYLKYSDEOLNCRGSPSPQNEWARAPKHIEDEIKQSEORQSRNQSTTTPYVTE 1080  
1021 LDTPINYLKYSDEOLNCRGSPSPQNEWARAPKHIEDEIKQSEORQSRNQSTTTPYVTE 1080  
1081 STDDKHLKFPQHFQGOECVSPYRSRGANGSETNRVGSNHGINQVNSQSLCOEDDEDYDDKP 1140  
1081 STDDKHLKFPQHFQGOECVSPYRSRGANGSETNRVGSNHGINQVNSQSLCOEDDEDYDDKP 1140  
1141 TNYISRYSEEBEHEEERTNYISIKYNEEKHVDOPIDYSLKYATDIPSSOKQGFSPSKS 1200  
1141 TNYISRYSEEBEHEEERTNYISIKYNEEKHVDOPIDYSLKYATDIPSSOKQGFSPSKS 1200  
1201 SSGQSKTEHMSSENSTSPSSNAKRONQLHPSSAQSRGQPOKAACTKVSSINQETIQ 1260  
1201 SSGQSKTEHMSSENSTSPSSNAKRONQLHPSSAQSRGQPOKAACTKVSSINQETIQ 1260  
1261 TYCVEDTPICFRSCSSLSLSAEDEIGCNQTTQPADSANTLOIAIEIKEI GTRSAEDPV 1320  
1261 TYCVEDTPICFRSCSSLSLSAEDEIGCNQTTQPADSANTLOIAIEIKEI GTRSAEDPV 1320  
1321 SEVPASVQHPRTKSRLOQSSLSSESARHKAVEFSSGAKSPSKSCAQTPKPEHYVOET 1380  
1321 SEVPASVQHPRTKSRLOQSSLSSESARHKAVEFSSGAKSPSKSCAQTPKPEHYVOET 1380  
1381 PLMFRCSTSVSLDSFESRSIASSVQSEPCSGMWSGIISPSLDPSQGTWPPRSKTKTP 1440  
1381 PLMFRCSTSVSLDSFESRSIASSVQSEPCSGMWSGIISPSLDPSQGTWPPRSKTKTP 1440  
1441 PPQTAQTKREVPKNAKPTAEKRESGPKQAAVNAVQVQLPDADTLLHFATESTPDGF 1500  
1441 PPQTAQTKREVPKNAKPTAEKRESGPKQAAVNAVQVQLPDADTLLHFATESTPDGF 1500  
1501 SCSSLSALSLEDEPTQKDVLRIMPVOENDNGNETSEQPKENENQKEAEKTI DSE 1560  
1501 SCSSLSALSLEDEPTQKDVLRIMPVOENDNGNETSEQPKENENQKEAEKTI DSE 1560  
1561 KOLLDDSDDDIIEILEECIISAMPTKSRKAKKPAQTASKLPPPVARKPSQLPVYKLLPS 1620  
1561 KOLLDDSDDDIIEILEECIISAMPTKSRKAKKPAQTASKLPPPVARKPSQLPVYKLLPS 1620  
1621 QNRLOPKHVSFTPGDDMPRVYCVETGPIINFSTATSLSLDTIESPNEIAGEGVRGGAQ 1680  
1621 QNRLOPKHVSFTPGDDMPRVYCVETGPIINFSTATSLSLDTIESPNEIAGEGVRGGAQ 1680  
1681 SGFEFKRDTIPTEGRSTDEAQGGKTSSTVPIPELDDNKAEEGDILAECINSAMPKGSHPK 1740  
1681 SGFEFKRDTIPTEGRSTDEAQGGKTSSTVPIPELDDNKAEEGDILAECINSAMPKGSHPK 1740  
1741 FRVKITMDVQQASASSAPNKNOLDGKKKTPSPVKPIPQNTTEYTRVRKNADSKNNLN 1800  
1741 FRVKITMDVQQASASSAPNKNOLDGKKKTPSPVKPIPQNTTEYTRVRKNADSKNNLN 1800  
1801 AERVFSDNKKONLKNNSKDFNDKLPNNEDRVGSPAFDPSPHYPTIEGTPYCFSRND 1860  
1801 AERVFSDNKKONLKNNSKDFNDKLPNNEDRVGSPAFDPSPHYPTIEGTPYCFSRND 1860  
1861 SLSSLDFFDDDDVLDLSREKAEKAKENKSEAKVTSHTELTNSQOSANKTOIAKQPINR 1920  
1861 SLSSLDFFDDDDVLDLSREKAEKAKENKSEAKVTSHTELTNSQOSANKTOIAKQPINR 1920  
1921 GQPKPILOKQSTPQSSKDIPDRGAATDEKLQNFAIENTPVCFSHNSLSLSDIDQENN 1980

Db 1921 GQPKPILOKQSTPQSSKDIPDRGAATDEKLQNFAIENTPVCFSHNSLSLSDIDQENN 1980  
QY 1981 KNEKEPIKETEPDPSOGEPSPKQASGYAPKSPHVEDTPVCFGRNSLSLSIDSEDDLQ 2040  
Db 1981 KNEKEPIKETEPDPSOGEPSPKQASGYAPKSPHVEDTPVCFGRNSLSLSIDSEDDLQ 2040  
QY 2041 ECITSSAMPKKKPSRLKGDNEKHSRNMGGIIGEDLTLDLKOIQRDPDSEHGLSPDSENF 2100  
Db 2041 ECITSSAMPKKKPSRLKGDNEKHSRNMGGIIGEDLTLDLKOIQRDPDSEHGLSPDSENF 2100  
QY 2101 WKAIQEGANSIVSSLHQAACLSRQASDSDSILSLKSGISLGSPPFHLTPQBEKPF 2160  
Db 2101 WKAIQEGANSIVSSLHQAACLSRQASDSDSILSLKSGISLGSPPFHLTPQBEKPF 2160  
QY 2161 SNKGPRILKPEKSTLTETTKIESESKIGKGVKYSKLITGKVRNSSEISGQWKQLOAN 2220  
Db 2161 SNKGPRILKPEKSTLTETTKIESESKIGKGVKYSKLITGKVRNSSEISGQWKQLOAN 2220  
QY 2221 MPISISGRTHIHPGVNRSSTSPVSKGPPKLTTPASKSPSEGTATTSPRCAKSPVKS 2280  
Db 2221 MPISISGRTHIHPGVNRSSTSPVSKGPPKLTTPASKSPSEGTATTSPRCAKSPVKS 2280  
QY 2281 ELSPVARQTSQIGGSKAPSRGSRDSTPSRPAQQLSRPIQSPGRNSISPGRNGISPPN 2340  
Db 2281 ELSPVARQTSQIGGSKAPSRGSRDSTPSRPAQQLSRPIQSPGRNSISPGRNGISPPN 2340  
QY 2341 KLSQLPRTSPSTASTKSSGSGKSVYTSRGRMSQOQLTKOTGLSKNASSIPSESEASK 2400  
Db 2341 KLSQLPRTSPSTASTKSSGSGKSVYTSRGRMSQOQLTKOTGLSKNASSIPSESEASK 2400  
QY 2401 LNMWNGNGANKVELSRMSSTKSSGSESDRSPRVLVROSTIKEAPSTLRRKLEESA 2460  
Db 2401 LNMWNGNGANKVELSRMSSTKSSGSESDRSPRVLVROSTIKEAPSTLRRKLEESA 2460  
QY 2461 SPESLSPSPSPASPTRSQAOTPVLSPLPMSISTHSSVOAGGWRKLPPLNSTIETNDG 2520  
Db 2461 SPESLSPSPSPASPTRSQAOTPVLSPLPMSISTHSSVOAGGWRKLPPLNSTIETNDG 2520  
QY 2521 RPAKHGDIAASHSPSRLPINRSGTWKREHSHSSLPVSTWRTTGGSSSILSASSES 2580  
Db 2521 RPAKHGDIAASHSPSRLPINRSGTWKREHSHSSLPVSTWRTTGGSSSILSASSES 2580  
QY 2581 SEKAKSEDEKHVNSISGOTKSKENQVSAKGTWAKIKENEFSPNTSQTQVSSGATNGAES 2640  
Db 2581 SEKAKSEDEKHVNSISGOTKSKENQVSAKGTWAKIKENEFSPNTSQTQVSSGATNGAES 2640  
QY 2641 KTLIYQMAPAVSKTEDVWVRIEDCPINNPRSGSPTCNTPPVIDSYSEKANPNIKSDKN 2700  
Db 2641 KTLIYQMAPAVSKTEDVWVRIEDCPINNPRSGSPTCNTPPVIDSYSEKANPNIKSDKN 2700  
QY 2701 QAKQNVGNGSVPMRTVGLNRLNSFIQVADAPDQKGTIKEPQONNPVSVSETNESSIVERT 2760  
Db 2701 QAKQNVGNGSVPMRTVGLNRLNSFIQVADAPDQKGTIKEPQONNPVSVSETNESSIVERT 2760  
QY 2761 PFSSSSSKHSSPSGTVAAVTPPNVNPSPRKSADSTSRPSQIPTPVNNNTKKRDSKT 2820  
Db 2761 PFSSSSSKHSSPSGTVAAVTPPNVNPSPRKSADSTSRPSQIPTPVNNNTKKRDSKT 2820  
QY 2821 DSTESSGTQSPKRHSGSYLVTSV 2843  
Db 2821 DSTESSGTQSPKRHSGSYLVTSV 2843

RESULT 5  
US-09-230-111C-30  
; Sequence 30, Application US/09230111C  
; Publication No. US20030203414A1  
; GENERAL INFORMATION:  
; APPLICANT: Sato, Taka-Aki  
; APPLICANT: Yanagisawa, Junn  
; TITLE OF INVENTION: COMPOUNDS THAT INHIBIT INTERACTION BETWEEN  
; TITLE OF INVENTION: SIGNAL-TRANSDUCING PROTEINS AND THE GLGF (PDZ/DHR)  
; TITLE OF INVENTION: DOMAIN AND USES THEREOF

FILE REFERENCE: 48962-A-PCT-US  
 ; CURRENT APPLICATION NUMBER: US  
 ; CURRENT FILING DATE: 1999-05-17  
 ; NUMBER OF SEQ ID NOS: 33  
 ; SOFTWARE: Patent in Ver. 2.1  
 ; SEQ ID NO 30  
 ; LENGTH: 2843  
 ; TYPE: PRT  
 ; ORGANISM: human  
 US-09-230-111C-30

Query Match 99.8%; Score 14539; DB 11; Length 2843;  
 Best Local Similarity 99.8%; Pred. No. 0;  
 Matches 2837; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy	1	MAAASYDQLLKQVEALXWENLNRLQELLEDNSNHLTKLEASNMKEVYLKOLQGSIEDAM	60
Db	1	MAAASYDQLLKQVEALXWENLNRLQELLEDNSNHLTKLEASNMKEVYLKOLQGSIEDAM	60
Qy	61	ASSGQIDLLERLKEINLDSSNFPQGVKLRKMSLRYSRGSGVSSRSRSGECPVPMGSPFR	120
Db	61	ASSGQIDLLERLKEINLDSSNFPQGVKLRKMSLRYSRGSGVSSRSRSGECPVPMGSPFR	120
Qy	121	RGFVNGRESGYLEELKERSLLADLDKEEKEKDYQAQLNLTKRIDSLPLTENFSL	180
Db	121	RGFVNGRESGYLEELKERSLLADLDKEEKEKDYQAQLNLTKRIDSLPLTENFSL	180
Qy	181	QTDVTRQLEYEARIIVAMEEQQLGTQDMKRAQRRIARIQIQIEKDILIRQLLOQAT	240
Db	181	QTDVTRQLEYEARIIVAMEEQQLGTQDMKRAQRRIARIQIQIEKDILIRQLLOQAT	240
Qy	241	EAERSSQNKHGTGSHDABRONEGGVGEINMATSGNGQGGTTRMDHETASVLSSTHSA	300
Db	241	EAERSSQNKHGTGSHDABRONEGGVGEINMATSGNGQGGTTRMDHETASVLSSTHSA	300
Qy	301	PRRLTSHLGTKEVYVLLSMLGTHDKDDMSRTLLAMSSQSDSCINRQSGCLPLLIQLL	360
Db	301	PRRLTSHLGTKEVYVLLSMLGTHDKDDMSRTLLAMSSQSDSCINRQSGCLPLLIQLL	360
Qy	361	HGNDKDSVLLNGSRGSKAARASAAALHNIHSQPDQKRRRIRVLHLLEQIRAYCETC	420
Db	361	HGNDKDSVLLNGSRGSKAARASAAALHNIHSQPDQKRRRIRVLHLLEQIRAYCETC	420
Qy	421	WEMQEAHEPGMDQKMPAPVEHOICPACVLMKLSFDEEHRHANNELGGQIAIELLO	480
Db	421	WEMQEAHEPGMDQKMPAPVEHOICPACVLMKLSFDEEHRHANNELGGQIAIELLO	480
Qy	481	VDCEMYGLTNDHYSITLRRYAGMALTNLTFGDVANKATLCSMKGCMRALVAQLKSEEDL	540
Db	481	VDCEMYGLTNDHYSITLRRYAGMALTNLTFGDVANKATLCSMKGCMRALVAQLKSEEDL	540
Qy	541	QOVIASVLRNLWRADVNSKTLREVGSVKALMECALEVVKESTLKSVALWLSAHTC	600
Db	541	QOVIASVLRNLWRADVNSKTLREVGSVKALMECALEVVKESTLKSVALWLSAHTC	600
Qy	601	ENKADICAVDGAFLAVGLTYRSQNTNLAIIESGGILRNVSLLIATNEDHRIQLRENN	660
Db	601	ENKADICAVDGAFLAVGLTYRSQNTNLAIIESGGILRNVSLLIATNEDHRIQLRENN	660
Qy	661	CIQTLLQHLKSHSLTIVSNACGLTWNLSARNPKDQALWDMGVMSLKNLIHSKHKMIAM	720
Db	661	CIQTLLQHLKSHSLTIVSNACGLTWNLSARNPKDQALWDMGVMSLKNLIHSKHKMIAM	720
Qy	721	GSAALRLNLMANPAKYKDANIMSPGSSLPVLRVKKOKALEAELDAQHLSETFDNIDLS	780
Db	721	GSAALRLNLMANPAKYKDANIMSPGSSLPVLRVKKOKALEAELDAQHLSETFDNIDLS	780
Qy	781	PKASHRSKQBHKSLQYGVYFDNRHDDNRDNFNFTGNMTVLSPYNTLTVLPSSSSRGS	840
Db	781	PKASHRSKQBHKSLQYGVYFDNRHDDNRDNFNFTGNMTVLSPYNTLTVLPSSSSRGS	840
Qy	841	LDSSRSKDRSLERERIGLGNYPATENPGTSSKRGQLISTTAAQIAKVMEVSAIHTS	900

Db	841	LDSSRSKDRSLERERIGLGNYPATENPGTSSKRGQLISTTAAQIAKVMEVSAIHTS	900
Qy	901	QEDRSSTGTTTELHCVTDERNALRRSSAAHSHNTNTFTKSENSNRTCSNMPYAKLEYKRSS	960
Db	901	QEDRSSTGTTTELHCVTDERNALRRSSAAHSHNTNTFTKSENSNRTCSNMPYAKLEYKRSS	960
Qy	961	NDLSNSVSSDGYGKRGQWKPSIESYSDDBSKFCYGYQADLAHKTSHANHMDNDGE	1020
Db	961	NDLSNSVSSDGYGKRGQWKPSIESYSDDBSKFCYGYQADLAHKTSHANHMDNDGE	1020
Qy	1021	LDTPINYSKYSDEQLNSGRQSPQNERWARPKHIIIEDEIKOSEBORGRNSTTYPVYTE	1080
Db	1021	LDTPINYSKYSDEQLNSGRQSPQNERWARPKHIIIEDEIKOSEBORGRNSTTYPVYTE	1080
Qy	1081	STDDKHLKFPQHFQGCVCVSPYRSGANGSETNRVGSNHGINQNVOSLCOEDDYEDDKP	1140
Db	1081	STDDKHLKFPQHFQGCVCVSPYRSGANGSETNRVGSNHGINQNVOSLCOEDDYEDDKP	1140
Qy	1141	TNYSERYSEEQHEEERPTNYSIKYNEEKRHVDQPIDYSLKYATDIPSSQKQSFSPKS	1200
Db	1141	TNYSERYSEEQHEEERPTNYSIKYNEEKRHVDQPIDYSLKYATDIPSSQKQSFSPKS	1200
Qy	1201	SSGQSKTEHSSSENTSTPSSNAKRONQLHPSAQSRSQPOKAATCKVSSINQETIQ	1260
Db	1201	SSGQSKTEHSSSENTSTPSSNAKRONQLHPSAQSRSQPOKAATCKVSSINQETIQ	1260
Qy	1261	TYCVEDTPICFRSCSSLSLSSAEDEIGCNOTTOEADSANTLQIAEIKETIGTSAEDPV	1320
Db	1261	TYCVEDTPICFRSCSSLSLSSAEDEIGCNOTTOEADSANTLQIAEIKETIGTSAEDPV	1320
Qy	1321	SEVPAVSQHPRTKSRLOGSSLSSESAPHKAVFSSGAKSPSKSGAQTTPKSPPHYQET	1380
Db	1321	SEVPAVSQHPRTKSRLOGSSLSSESAPHKAVFSSGAKSPSKSGAQTTPKSPPHYQET	1380
Qy	1381	PLMFSRCTSVSLDSFBSRSIASSVQSEPCSGMVGIIISPSDLPSDQTMPPRSRSTPP	1440
Db	1381	PLMFSRCTSVSLDSFBSRSIASSVQSEPCSGMVGIIISPSDLPSDQTMPPRSRSTPP	1440
Qy	1441	PPPTAQTKREVVPKNKAPTAEKRESGPKQAANAQVQVLPDADTLILHFAFESTPDGF	1500
Db	1441	PPPTAQTKREVVPKNKAPTAEKRESGPKQAANAQVQVLPDADTLILHFAFESTPDGF	1500
Qy	1501	SCSSLSLSLSDLEPFIQKDVLRIMPPVOENDNGNETESQPKESNENQKEAKETIDSE	1560
Db	1501	SCSSLSLSLSDLEPFIQKDVLRIMPPVOENDNGNETESQPKESNENQKEAKETIDSE	1560
Qy	1561	KDLLDSDDDDIIELEECIIISAMPTKSRKAKPAQATASKLPPPVARKPSQLPVYKLLPS	1620
Db	1561	KDLLDSDDDDIIELEECIIISAMPTKSRKAKPAQATASKLPPPVARKPSQLPVYKLLPS	1620
Qy	1621	QNRLOPKHVSFTQDDMPRVYVEGTPINPSTATSLDITSPPNELAAAGEVGRGAQ	1680
Db	1621	QNRLOPKHVSFTQDDMPRVYVEGTPINPSTATSLDITSPPNELAAAGEVGRGAQ	1680
Qy	1681	SGEFKEDDTIIPTEGRSTDEAOGGTSSTVTIPELDDNKAEGDIIAECINSAMPKSKHHP	1740
Db	1681	SGEFKEDDTIIPTEGRSTDEAOGGTSSTVTIPELDDNKAEGDIIAECINSAMPKSKHHP	1740
Qy	1741	FRVKIMQVQQAASASSAPNQLDGKKKPTSPVPIFQNTYRTRVRKNADSKNIN	1800
Db	1741	FRVKIMQVQQAASASSAPNQLDGKKKPTSPVPIFQNTYRTRVRKNADSKNIN	1800
Qy	1801	AERFSDNDSKKQNLKXNSKDFNDKLPNNEDRVGSAFDSPPHYTPIEGTYPYCFSRND	1860
Db	1801	AERFSDNDSKKQNLKXNSKDFNDKLPNNEDRVGSAFDSPPHYTPIEGTYPYCFSRND	1860
Qy	1861	SLSLDLDLDDVDLSRKAELRKAENKESAKVTSHTELTSNQOQSANKTQAIAKQPINR	1920
Db	1861	SLSLDLDLDDVDLSRKAELRKAENKESAKVTSHTELTSNQOQSANKTQAIAKQPINR	1920
Qy	1921	GQPKPIIKQSTFPQSSKQIPDRGAATDEKLQNFATIENTPVCFSHNSLSLSDIDENN	1980

```
QY 1981 NKENPPIKETPPDQGGPSPKQASGYAPKSPHVEDTPVCFSRNSSLSSLSIDSEDLQ 2040
Db 1981 NKENPPIKETPPDQGGPSPKQASGYAPKSPHVEDTPVCFSRNSSLSSLSIDSEDLQ 2040
QY 2041 ECISAMPKKKPSRLKGDNEKHSRNNGGILGEDLTLDLKDIOQPDSEHGLSPDSENF 2100
Db 2041 ECISAMPKKKPSRLKGDNEKHSRNNGGILGEDLTLDLKDIOQPDSEHGLSPDSENF 2100
QY 2101 WKAIQEGANSIVSSLHQAAAACLSRQASDSDSLSLKSGISLGSPPHLPDQEEPPFT 2160
Db 2101 WKAIQEGANSIVSSLHQAAAACLSRQASDSDSLSLKSGISLGSPPHLPDQEEPPFT 2160
QY 2161 SNKGRPIILKPEKSTLETKKIESESGIKGGKKVYKSLITKVRNSNBSISQOMKQLOAN 2220
Db 2161 SNKGRPIILKPEKSTLETKKIESESGIKGGKKVYKSLITKVRNSNBSISQOMKQLOAN 2220
QY 2221 MPSISGRGTMHIIPGVNRNSSSTSPVSKGPPKTPASKSPSEGOATATSPRGAKPSVKS 2280
Db 2221 MPSISGRGTMHIIPGVNRNSSSTSPVSKGPPKTPASKSPSEGOATATSPRGAKPSVKS 2280
QY 2281 ELSPVARQTSQIGGSKAPSRGSRDSTSPRAQOPLSRPQSPGRNBSISPRNGISPPN 2340
Db 2281 ELSPVARQTSQIGGSKAPSRGSRDSTSPRAQOPLSRPQSPGRNBSISPRNGISPPN 2340
QY 2341 KLSQLPRTSSPSTASTKSGGSKMSYTPSGRQMSQONLTQGTGLSKNASSIPRSESASK 2400
Db 2341 KLSQLPRTSSPSTASTKSGGSKMSYTPSGRQMSQONLTQGTGLSKNASSIPRSESASK 2400
QY 2401 LNQMNGNGANKKVELSRMSSTKSSGSESDRSERPVLVRQFTFIKEAPSPTRLRKLSESA 2460
Db 2401 LNQMNGNGANKKVELSRMSSTKSSGSESDRSERPVLVRQFTFIKEAPSPTRLRKLSESA 2460
QY 2461 SPESLSPSRPASPTSCAQOFTVLSPLPDMSLSTHSSVQAGWRKLPNLSPTIEYNDG 2520
Db 2461 SPESLSPSRPASPTSCAQOFTVLSPLPDMSLSTHSSVQAGWRKLPNLSPTIEYNDG 2520
QY 2521 RPAKHEDLARSHSESPSRIPINRSCTWKREHSKSSSILPRVSTWRTGSSSSILSASSES 2580
Db 2521 RPAKHEDLARSHSESPSRIPINRSCTWKREHSKSSSILPRVSTWRTGSSSSILSASSES 2580
QY 2581 SEKAKSEDEKHVNSISGTKQKENVASAGTWKIKENEFPPTNSTQTSVSSGATNGAES 2640
Db 2581 SEKAKSEDEKHVNSISGTKQKENVASAGTWKIKENEFPPTNSTQTSVSSGATNGAES 2640
QY 2641 KTLIIQMAFPAVSKTDEWVRVEDCPINNPRSGRSPGTGNTPPVIDSVSEKAMPNIKSDKN 2700
Db 2641 KTLIIQMAFPAVSKTDEWVRVEDCPINNPRSGRSPGTGNTPPVIDSVSEKAMPNIKSDKN 2700
QY 2701 QAKQNVGSGVPMRTVGLNRLNSFIQVADPQKGTETIKPGQNNPVPVSETNESSIVERT 2760
Db 2701 QAKQNVGSGVPMRTVGLNRLNSFIQVADPQKGTETIKPGQNNPVPVSETNESSIVERT 2760
QY 2761 PFSSSSSKHSSPSTVAARVTPFNYPNPSPRKSSADSTSARPSQIPTPVNNTKKRDSKT 2820
Db 2761 PFSSSSSKHSSPSTVAARVTPFNYPNPSPRKSSADSTSARPSQIPTPVNNTKKRDSKT 2820
QY 2821 DSTESSGTSQPRKHSGSYLVTSV 2843
Db 2821 DSTESSGTSQPRKHSGSYLVTSV 2843
```

## RESULT 6

```
US-10-092-138-30
; Sequence 30, Application US/10092138
; Publication No. US20030170723A1
; GENERAL INFORMATION:
; APPLICANT: Sato, Taka-Aki
; TITLE OF INVENTION: METHOD OF PREPARING A PROTEIN ARRAY BASED ON
; TITLE OF INVENTION: BIOCHEMICAL PROTEIN-PROTEIN INTERACTION
; FILE REFERENCE: 65823/JPW/PT
; CURRENT APPLICATION NUMBER: US/10/092,138
; CURRENT FILING DATE: 2002-09-06
```

```
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 30
; LENGTH: 2843
; TYPE: PRT
; ORGANISM: human
US-10-092-138-30
```

```
Query Match 99.8%; Score 14539; DB 14; Length 2843;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2837; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 MAASVDQLLKQVEALKMENSILRQLEDNSNHLTKLETEASNKKEVLKOLQSGIDEAM 60
Db 1 MAASVDQLLKQVEALKMENSILRQLEDNSNHLTKLETEASNKKEVLKOLQSGIDEAM 60
QY 61 ASSGQIDLLERLKLNDSSNFPFVKLRKMSLRYSYSGREGSVSSRSGECPVPMGSPFR 120
Db 61 ASSGQIDLLERLKLNDSSNFPFVKLRKMSLRYSYSGREGSVSSRSGECPVPMGSPFR 120
QY 121 RGFVNGSRESTGYLEBELEKERSILLADLKEEKEKDWYAOQLNLTKRDLPLTENFSL 180
Db 121 RGFVNGSRESTGYLEBELEKERSILLADLKEEKEKDWYAOQLNLTKRDLPLTENFSL 180
QY 181 QTDWTRQLEVEARQIRVAMEEOLGTCDQMEKRAQRRIARIQOIEKDILRIROLLOQAT 240
Db 181 QTDWTRQLEVEARQIRVAMEEOLGTCDQMEKRAQRRIARIQOIEKDILRIROLLOQAT 240
QY 241 EAERSSONKHETGSHDAERQNEGGVGEINMATSGNGQSGSTTRMDHETASVLSSTHSA 300
Db 241 EAERSSONKHETGSHDAERQNEGGVGEINMATSGNGQSGSTTRMDHETASVLSSTHSA 300
QY 301 PRRLTSHLGTKEWYVSLLSMLGTHDKDDMSRTLLAMSSQSDSCISMRQSGCLPLLIQLL 360
Db 301 PRRLTSHLGTKEWYVSLLSMLGTHDKDDMSRTLLAMSSQSDSCISMRQSGCLPLLIQLL 360
QY 361 HGNDKDSVLLGNGRSGKEARASAAALHNIHSQDPDKRGRREIRVHLHLEQIRAYCETC 420
Db 361 HGNDKDSVLLGNGRSGKEARASAAALHNIHSQDPDKRGRREIRVHLHLEQIRAYCETC 420
QY 421 WEWQEAHEPGMDQKNMPAPAVEHQICPAVCVLMKLSFDEEHHANNEIGLOAIAELLQ 480
Db 421 WEWQEAHEPGMDQKNMPAPAVEHQICPAVCVLMKLSFDEEHHANNEIGLOAIAELLQ 480
QY 481 VDCMYGLTNDHYSITLRRYAGMALTNLTFDGVANKATLCSMGKGCVRALVAQLKSSSEDL 540
Db 481 VDCMYGLTNDHYSITLRRYAGMALTNLTFDGVANKATLCSMGKGCVRALVAQLKSSSEDL 540
QY 541 QQVITASVLRNLSWRADVNSKKTIREVGSVKALMECALFVKKESTLKSVALNLSAHCT 600
Db 541 QQVITASVLRNLSWRADVNSKKTIREVGSVKALMECALFVKKESTLKSVALNLSAHCT 600
QY 601 ENKADICAVDGAFLAVGLTYRSQNTLAIIESGGILRNVSLSLIATNEDHQILRENN 660
Db 601 ENKADICAVDGAFLAVGLTYRSQNTLAIIESGGILRNVSLSLIATNEDHQILRENN 660
QY 661 CLQTLLOHLKSHSLTIVSNACGTLNLSARNPKDQBALMDMGAVSMLKNLIHSHKXMIAM 720
Db 661 CLQTLLOHLKSHSLTIVSNACGTLNLSARNPKDQBALMDMGAVSMLKNLIHSHKXMIAM 720
QY 721 GSAALLENLMANEPKAKYKDNATMSPGSSLPSEVRKQKALEAELDAQHLSSETDNDLS 780
Db 721 GSAALLENLMANEPKAKYKDNATMSPGSSLPSEVRKQKALEAELDAQHLSSETDNDLS 780
QY 781 PKASHRSKQKHQSLYGDYVDFDNRHDDNRSDNFNTGNMTVLSPLYLNTTVLPSSSSSRGS 840
Db 781 PKASHRSKQKHQSLYGDYVDFDNRHDDNRSDNFNTGNMTVLSPLYLNTTVLPSSSSSRGS 840
QY 841 LDSRSSEKRSLEIRERIGLGNTHPATENPGTSSKGLQISTTAAQIAKMWESVAIHTS 900
Db 841 LDSRSSEKRSLEIRERIGLGNTHPATENPGTSSKGLQISTTAAQIAKMWESVAIHTS 900
QY 901 QEDRSSGSGTTELHCVTDERNALRRSSAAHSTHNTYNTFTKSENSNRITCSMEYAKLEYKRSS 960
```

Db 901 QEDRSSGTTTLCVTDERNALRRSSAAHTSHNTYNTFKSENRR:CSMPYAKLEYKRS 960  
Qy 961 NDSLSYSSSDGKGKQKMPKPS:IESXSEDDSEKFCYGOYPADLAHKIHSANHMDNDGE 1020  
Db 961 NDSLSYSSSDGKGKQKMPKPS:IESXSEDDSEKFCYGOYPADLAHKIHSANHMDNDGE 1020  
Qy 1021 LDTPIYNSLYKSDPOLNSGQSQSONERWAPXHI:IEDEIKQSEQRQSNQSTTYPVYTE 1080  
Db 1021 LDTPIYNSLYKSDPOLNSGQSQSONERWAPXHI:IEDEIKQSEQRQSNQSTTYPVYTE 1080  
Qy 1081 STDDKHLKFQPHFGQBCVSPYRSRANGSETNRVGSNGHINQVQSICQEDDYEDDKP 1140  
Db 1081 STDDKHLKFQPHFGQBCVSPYRSRANGSETNRVGSNGHINQVQSICQEDDYEDDKP 1140  
Qy 1141 TNSYRYSEBEOHEEBERPTNYISIKNEEKRVHDQPIDYSLKYATDIPSSQKQSPFSKS 1200  
Db 1141 TNSYRYSEBEOHEEBERPTNYISIKNEEKRVHDQPIDYSLKYATDIPSSQKQSPFSKS 1200  
Qy 1201 SSGQSKTEHWSSENSTSTPSSNAKRONLHPSSAQSRGQPKAATCKVSSINQETIQ 1260  
Db 1201 SSGQSKTEHWSSENSTSTPSSNAKRONLHPSSAQSRGQPKAATCKVSSINQETIQ 1260  
Qy 1261 TYCVEDTPICFRCSLSLSSLSAEDEIGCQNTTQEADSAANTLQIAIEIKKIGTRSAEDPV 1320  
Db 1261 TYCVEDTPICFRCSLSLSSLSAEDEIGCQNTTQEADSAANTLQIAIEIKKIGTRSAEDPV 1320  
Qy 1321 SEVPAVQCHRTKSSRLQGSLSAESEARHKA:VFSSGAKSPSKGACTPKSPPEHYVQET 1380  
Db 1321 SEVPAVQCHRTKSSRLQGSLSAESEARHKA:VFSSGAKSPSKGACTPKSPPEHYVQET 1380  
Qy 1381 PLMFSRCTSVSSLDSPESRSIASSVQSEPCSGMVSGI:ISPSDLDPFGQTMPPSRSKTTP 1440  
Db 1381 PLMFSRCTSVSSLDSPESRSIASSVQSEPCSGMVSGI:ISPSDLDPFGQTMPPSRSKTTP 1440  
Qy 1441 PPQTAQTKREVKNKAPTAKESGPKQAANVAVQVQVLPDADTLHFATESPDGF 1500  
Db 1441 PPQTAQTKREVKNKAPTAKESGPKQAANVAVQVQVLPDADTLHFATESPDGF 1500  
Qy 1501 SCSSLSALSILDEPFIQKVELRIMPVQENDNGNETESQPKESNENQKAEKTI:DS 1560  
Db 1501 SCSSLSALSILDEPFIQKVELRIMPVQENDNGNETESQPKESNENQKAEKTI:DS 1560  
Qy 1561 KDLLDDSDDDDI:ELIEECIIISAMPTKSRKAKPAOTASKLPPPVARKPSQLPVYKLLPS 1620  
Db 1561 KDLLDDSDDDDI:ELIEECIIISAMPTKSRKAKPAOTASKLPPPVARKPSQLPVYKLLPS 1620  
Qy 1621 QNRLOPKHVSFTPGDDMPRVYCVGTPINFSTATSLSDLTIESPPNELAAGBVGGAQ 1680  
Db 1621 QNRLOPKHVSFTPGDDMPRVYCVGTPINFSTATSLSDLTIESPPNELAAGBVGGAQ 1680  
Qy 1681 SGFEKEDTITPTEGRSTDEAGGKTSSVTI:PELDDNKABEGDILABEINCSAMPKGSHPK 1740  
Db 1681 SGFEKEDTITPTEGRSTDEAGGKTSSVTI:PELDDNKABEGDILABEINCSAMPKGSHPK 1740  
Qy 1741 FRVKIMQVQQAASASSAPNKQDGGKKKPTSPVKPI:PONTEYTRVRKNADSNNLN 1800  
Db 1741 FRVKIMQVQQAASASSAPNKQDGGKKKPTSPVKPI:PONTEYTRVRKNADSNNLN 1800  
Qy 1801 AERFSDNKKQKQNLKNSKDFNDKLPNNEDVRGSPAFDPSPHYTPIEGTPYCFSRND 1860  
Db 1801 AERFSDNKKQKQNLKNSKDFNDKLPNNEDVRGSPAFDPSPHYTPIEGTPYCFSRND 1860  
Qy 1861 SLSLDDDDDDVLSRKAELRKAENKESAKVTSHT:ELTSNOQANKTQATKAPINR 1920  
Db 1861 SLSLDDDDDDVLSRKAELRKAENKESAKVTSHT:ELTSNOQANKTQATKAPINR 1920  
Qy 1921 GQPKPILOKQSTFPQSSKIDPDRGAATDEKLQNFALIENTPVCFSHNSLSLSLSDIDQENN 1980  
Db 1921 GQPKPILOKQSTFPQSSKIDPDRGAATDEKLQNFALIENTPVCFSHNSLSLSLSDIDQENN 1980  
Qy 1981 NKENPIKETEPDPSQGEPSKQASGYAPKSFHVEDTPVCFSRNSLSLSLSDIDSEDDLLQ 2040

Db 1981 NKENPIKETEPDPSQGEPSKQASGYAPKSFHVEDTPVCFSRNSLSLSLSDIDSEDDLLQ 2040  
Qy 2041 ECISAMPKKKKPRLKGDNEKXHSFRNMGGLGSDLTLDLKD:IORPSEHGLSPDSNFD 2100  
Db 2041 ECISAMPKKKKPRLKGDNEKXHSFRNMGGLGSDLTLDLKD:IORPSEHGLSPDSNFD 2100  
Qy 2101 WKATQEGANSIVSSIHQAAAAACLSROASDSDSILSLKSGISLGSPPHLP:TPDOEEKPFT 2160  
Db 2101 WKATQEGANSIVSSIHQAAAAACLSROASDSDSILSLKSGISLGSPPHLP:TPDOEEKPFT 2160  
Qy 2161 SNKGPRILKPEKSTLETKTIESESKGKIGKGVKYSILITGKVRNSSEI:SQMKQPLQAN 2220  
Db 2161 SNKGPRILKPEKSTLETKTIESESKGKIGKGVKYSILITGKVRNSSEI:SQMKQPLQAN 2220  
Qy 2221 MPSISRGFTMIHIGVNSSSSTSPVSKGPPPLKTPASKSPSEGTATTSPRGA:KPSVKS 2280  
Db 2221 MPSISRGFTMIHIGVNSSSSTSPVSKGPPPLKTPASKSPSEGTATTSPRGA:KPSVKS 2280  
Qy 2281 ELSPVARQTSQIGGSKAPSRSGSRDSTPRPAQOPLSRPIQSPGRNSI:SPGRNGISPPN 2340  
Db 2281 ELSPVARQTSQIGGSKAPSRSGSRDSTPRPAQOPLSRPIQSPGRNSI:SPGRNGISPPN 2340  
Qy 2341 KLSQLPRTSPSTASTKSSSGCKMSYTSFGROMSQOQLTKOTGLSKNASSI:PRASESASKG 2400  
Db 2341 KLSQLPRTSPSTASTKSSSGCKMSYTSFGROMSQOQLTKOTGLSKNASSI:PRASESASKG 2400  
Qy 2401 LNQMNGANGANKVELSRMSSTKSSGSEDRSERPVLVRQSTFTIKEAP:SPTLRKLBSA 2460  
Db 2401 LNQMNGANGANKVELSRMSSTKSSGSEDRSERPVLVRQSTFTIKEAP:SPTLRKLBSA 2460  
Qy 2461 SFESLSPSSRPASPTRSOAQTPLVSLPMSLSTHSSVQAGGWRKLP:PKMLSPITIEYNDG 2520  
Db 2461 SFESLSPSSRPASPTRSOAQTPLVSLPMSLSTHSSVQAGGWRKLP:PKMLSPITIEYNDG 2520  
Qy 2521 RPAKHDIARSHSSPRLPINRSGTWKRHSKHSSSLPRVSTWRTGSSS:ILSASSES 2580  
Db 2521 RPAKHDIARSHSSPRLPINRSGTWKRHSKHSSSLPRVSTWRTGSSS:ILSASSES 2580  
Qy 2581 SEKAKEDEKHVNSISGTKQSKENQVSAKGTWRIK:ENEFSPNTNSTQTVSSGATNGAES 2640  
Db 2581 SEKAKEDEKHVNSISGTKQSKENQVSAKGTWRIK:ENEFSPNTNSTQTVSSGATNGAES 2640  
Qy 2641 KTLIYQWAPAVSKTEDVWVRIEDCPINNPRSGRSPGTGNTPPVIDSV:SEKANENIKDSKN 2700  
Db 2641 KTLIYQWAPAVSKTEDVWVRIEDCPINNPRSGRSPGTGNTPPVIDSV:SEKANENIKDSKN 2700  
Qy 2701 QAKQNVGNGSVPMRTVGLNRLNSFIOVDAPDQXGTEIKPGQNNPVPV:SETNESSIYERT 2760  
Db 2701 QAKQNVGNGSVPMRTVGLNRLNSFIOVDAPDQXGTEIKPGQNNPVPV:SETNESSIYERT 2760  
Qy 2761 PFSSSSSKHSSPSTGTVAAARVTPFNYPNPSPKSADSTSA:RPSQIPTPVANNNTKRSK 2820  
Db 2761 PFSSSSSKHSSPSTGTVAAARVTPFNYPNPSPKSADSTSA:RPSQIPTPVANNNTKRSK 2820  
Qy 2821 DSTSSGTSQSPKRHSGLVTSV 2843  
Db 2821 DSTSSGTSQSPKRHSGLVTSV 2843

## RESULT 7

US-10-267-502-370  
; Sequence 370, Application US/10267502  
; Publication No. US20040071700A1  
; GENERAL INFORMATION:  
; APPLICANT: Kim, Jaeseob  
; APPLICANT: Galant, Ron  
; TITLE OF INVENTION: Obesity Linked Genes  
; FILE REFERENCE: LSD-07416  
; CURRENT APPLICATION NUMBER: US/10/267,502  
; CURRENT FILING DATE: 2003-01-27  
; NUMBER OF SEQ ID NOS: 439  
; SOFTWARE: Patent in version 3.2  
; SEQ ID NO 370

; LENGTH: 2844  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-267-502-370  
  
Query Match  
Best Local Similarity 99.4%; Score 14484.5; DB 12; Length 2844;  
Matches 2830; Conservative 3; Mismatches 10; Indels 1; Gaps 1;  
  
QY 1 MAAASYDQLLKQVEALKWENSLRQLEDNSNHLTKLETEASNKKVYLKQLQGSIEDEAM 60  
DB 1 MAAASYDQLLKQVEALKWENSLRQLEDNSNHLTKLETEASNKKVYLKQLQGSIEDEAM 60  
QY 61 ASSGOIDLLERLKEALNLDSSNPPGVKLRSKMSLRSYSGREGSVSRSGCECPVPMGSPPR 120  
DB 61 ASSGOIDLLERLKEALNLDSSNPPGVKLRSKMSLRSYSGREGSVSRSGCECPVPMGSPPR 120  
QY 121 RGFVNGSRSTGYLBELEKERSLILLADLDKBEKEDWYQAQNLTKRIDSPLTFENFSL 180  
DB 121 RGFVNGSRSTGYLBELEKERSLILLADLDKBEKEDWYQAQNLTKRIDSPLTFENFSL 180  
QY 181 QDMLTRRQLEYEARQIRVAMEEQLGTCODMEXRAORRIARICQIEKDILIRLQLOSAT 240  
DB 181 QDMLTRRQLEYEARQIRVAMEEQLGTCODMEXRAORRIARICQIEKDILIRLQLOSAT 240  
QY 241 EAERSSONKHETGSHDAERONEGGVGEINMATSGNGOGSTTRMDHETASVLSSSSTHSA 300  
DB 241 EAERSSONKHETGSHDAERONEGGVGEINMATSGNGOGSTTRMDHETASVLSSSSTHSA 300  
QY 301 PRLTSLHGTKVMYVLSLMSLGTDXDMSSTLLAMSSQDSCISMRCQSGCLPLLIQLL 360  
DB 301 PRLTSLHGTKVMYVLSLMSLGTDXDMSSTLLAMSSQDSCISMRCQSGCLPLLIQLL 360  
QY 361 HGNDDKSVLLGNRSKSKARASAAALHNIHISQPDCKGGRREIRVLHLQIRAYCETC 420  
DB 361 HGNDDKSVLLGNRSKSKARASAAALHNIHISQPDCKGGRREIRVLHLQIRAYCETC 420  
QY 421 WQQAHEFGMDQDNMPAPVEHQICPAVCVLMKLSFDEEHRHAMELGGIQAIAELLQ 480  
DB 421 WQQAHEFGMDQDNMPAPVEHQICPAVCVLMKLSFDEEHRHAMELGGIQAIAELLQ 480  
QY 481 VDCMYGLTNDHYSITLRYAGMALTNLTFGVDVANKATILCSMKGCWRALVAQLKSESD 540  
DB 481 VDCMYGLTNDHYSITLRYAGMALTNLTFGVDVANKATILCSMKGCWRALVAQLKSESD 540  
QY 541 QQVIASVLNLSWRADVNSKKTILREVGSVKALMECALEVKKESTLKSLSALWNLSAHCT 600  
DB 541 QQVIASVLNLSWRADVNSKKTILREVGSVKALMECALEVKKESTLKSLSALWNLSAHCT 600  
QY 601 ENKADICAVDGAFLVGLTYRSQNTLAIIESGGILRNVSLSIATNEDHRQILRENN 660  
DB 601 ENKADICAVDGAFLVGLTYRSQNTLAIIESGGILRNVSLSIATNEDHRQILRENN 660  
QY 661 CLQTLLOHLKSHSLTIVSNACGTLWNLARNPKDQALWDMGAVSMKLNLIHSHKMTIAM 720  
DB 661 CLQTLLOHLKSHSLTIVSNACGTLWNLARNPKDQALWDMGAVSMKLNLIHSHKMTIAM 720  
QY 721 GSAALRLNLMANRPARYKDANIMSPGSLPSLFVRKQKALEAELDAQHLSETFDNIDLS 780  
DB 721 GSAALRLNLMANRPARYKDANIMSPGSLPSLFVRKQKALEAELDAQHLSETFDNIDLS 780  
QY 781 PKASHRSKQKHOSLYGDYVDFNTRHDDNRSNFTNGNTVLSPLYNTTIVLPSSSSSRGS 840  
DB 781 PKASHRSKQKHOSLYGDYVDFNTRHDDNRSNFTNGNTVLSPLYNTTIVLPSSSSSRGS 840  
QY 841 LDSRSSEKORSLEERERGI GLGNYPHATENPGTSSKRGLOI STTAAQIAKMEEVSAIHTS 900  
DB 841 LDSRSSEKORSLEERERGI GLGNYPHATENPGTSSKRGLOI STTAAQIAKMEEVSAIHTS 900  
QY 901 QEDRSSGSTTELCVTDERNALRRSAAHTNSNTYNTFTKSENRTCSMPYAKLEYKRSS 960  
DB 901 QEDRSSGSTTELCVTDERNALRRSAAHTNSNTYNTFTKSENRTCSMPYAKLEYKRSS 960

QY 961 NDSLNSVSSDDGYGKRGQMKPSIESYSEDDSKFCSYGQYPADLAHKAHSAHMDNDGE 1020  
DB 961 NDSLNSVSSDDGYGKRGQMKPSIESYSEDDSKFCSYGQYPADLAHKAHSAHMDNDGE 1020  
QY 1021 LDTPIYSLKYSDQNLNQRSPSONERWARPKHIIIEDEIKQSEQRQSRNQSTTYPVYTE 1080  
DB 1021 LDTPIYSLKYSDQNLNQRSPSONERWARPKHIIIEDEIKQSEQRQSRNQSTTYPVYTE 1080  
QY 1081 STDKHLKFPQHGQOECVSPYRSRGANGSETNRVCSNHCINQNVQSICQOEDDYEDDKP 1140  
DB 1081 STDKHLKFPQHGQOECVSPYRSRGANGSETNRVCSNHCINQNVQSICQOEDDYEDDKP 1140  
QY 1141 TNYSERYSBEEQHEEERPTNYSIKYNEEKRVHQDQIDYSLKYATDIPSSQKQSFSPSKS 1200  
DB 1141 TNYSERYSBEEQHEEERPTNYSIKYNEEKRVHQDQIDYSLKYATDIPSSQKQSFSPSKS 1200  
QY 1201 SSGOSSKTEHMSSENSTSTPSSNAKQNLHPSSAQSRSGQGPQKAATCKVSSINQETIQ 1260  
DB 1201 SSGOSSKTEHMSSENSTSTPSSNAKQNLHPSSAQSRSGQGPQKAATCKVSSINQETIQ 1260  
QY 1261 TYCVEDTPICFSCSSISLSSAEDIEGNCQTTQEADSAANTLOIAETKEKIGTRSAEDPV 1320  
DB 1261 TYCVEDTPICFSCSSISLSSAEDIEGNCQTTQEADSAANTLOIAETKEKIGTRSAEDPV 1320  
QY 1321 SEVP-AVSOHPRTKSRLOGSSLSSESARHAKAVEPSSGAKSPSKSGAQTTPKSPPEHYVQE 1379  
DB 1321 SEVPSSVHSTLETKSSRLQSSLSSESARHAKAVEPSSGAKSPSKSGAQTTPKSPPEHYVQE 1380  
QY 1380 TPLMFSCTSVSLDSFESRSIASSVQSEPCSGWAGIISPSDLPSPGQTMPPSRKTP 1439  
DB 1381 TPLMFSCTSVSLDSFESRSIASSVQSEPCSGWAGIISPSDLPSPGQTMPPSRKTP 1440  
QY 1440 PPPQTAQTKREVPKNAPTAEKRESGPKQAAVNAAVQRVQVLPDADTLHLHFAESTFPG 1499  
DB 1441 PPPQTAQTKREVPKNAPTAEKRESGPKQAAVNAAVQRVQVLPDADTLHLHFAESTFPG 1500  
QY 1500 FSCSSLSALSALSBEPFTQKQVELRIMPVOENDNGNTESEQPKESNENQKAEKTIDS 1559  
DB 1501 FSCSSLSALSALSBEPFTQKQVELRIMPVOENDNGNTESEQPKESNENQKAEKTIDS 1560  
QY 1560 EKDLDDSDDDDEILIEECIIISAMPTKSSRKAKKPAOTASKLPPPVARKESQLPVYKLLP 1619  
DB 1561 EKDLDDSDDDDEILIEECIIISAMPTKSSRKAKKPAOTASKLPPPVARKESQLPVYKLLP 1620  
QY 1620 SQNRLOPKHVSFTPGDDMPRVYCVGCTPIINFSTATSLDITIESPPNBLAAGEVGRGA 1679  
DB 1621 SQNRLOPKHVSFTPGDDMPRVYCVGCTPIINFSTATSLDITIESPPNBLAAGEVGRGA 1680  
QY 1680 QSGPEKEDTIPTEGRSTDEAQGGKTSVTIPELDDNKAEEGDILAECSINAMPKGSKHK 1739  
DB 1681 QSGPEKEDTIPTEGRSTDEAQGGKTSVTIPELDDNKAEEGDILAECSINAMPKGSKHK 1740  
QY 1740 PFRVYKTMDOVQAASASSAPNKNQLDGKKKXPTSPVKPIPQNTYRTRVRKNADSKNKL 1799  
DB 1741 PFRVYKTMDOVQAASASSAPNKNQLDGKKKXPTSPVKPIPQNTYRTRVRKNADSKNKL 1800  
QY 1800 NAERVFSNDKSKQNLKNNSKDFNDKLPNNEDRVGSPAFDSPHHYTPTEGTPYCFSRN 1859  
DB 1801 NAERVFSNDKSKQNLKNNSKDFNDKLPNNEDRVGSPAFDSPHHYTPTEGTPYCFSRN 1860  
QY 1860 DLSLSLDFDDDDVLSREKAEKAKENKESAKVTSHTELTNSQOSANKTQAIKQPIN 1919  
DB 1861 DLSLSLDFDDDDVLSREKAEKAKENKESAKVTSHTELTNSQOSANKTQAIKQPIN 1920  
QY 1920 RGQPKPILOKQSTPPQSSKOIPORGAATDEKQLQNFALIENTPVCFSHNSLSLSLSDIDQEN 1979  
DB 1921 RGQPKPILOKQSTPPQSSKOIPORGAATDEKQLQNFALIENTPVCFSHNSLSLSLSDIDQEN 1980  
QY 1980 NKNENEPKETEPPDSOGEPKPOASGVAPKSFHVEDTPVCFSRNSSLSLSIDSEDDL 2039  
DB 1991 NKNENEPKETEPPDSOGEPKPOASGVAPKSFHVEDTPVCFSRNSSLSLSIDSEDDL 2040  
QY 2040 QECISSAMPKKKPKSRLKGDNKHSPRNMGGIIGEDLTLDKDIQRPDSEHGLSPDSENF 2099



Db 2041 QECISSAMPKKKSRLLKGDNEKSPNNMGILGEDJTLDDKQIQRDPDSEHGSLSPDSENF 2100  
Qy 2100 DWKAIQEGANGIVSLHQAAAACLSRQASDSDSIILSLKSGISLGSPFHLTPDQBEKPF 2159  
Db 2101 DWKAIQEGANGIVSLHQAAAACLSRQASDSDSIILSLKSGISLGSPFHLTPDQBEKPF 2160  
Qy 2160 TSNKGPRILKPEKSTLETKIYESGKIGKKVYKSLITGKVRNSEISGQMKPQLOA 2219  
Db 2161 TSNKGPRILKPEKSTLETKIYESGKIGKKVYKSLITGKVRNSEISGQMKPQLOA 2220  
Qy 2220 NMPISIRGRTHIHPGVNRSSTSPVSKGPPPLKTPASKSPGEGQTATTSRGAKEPSVK 2279  
Db 2221 NMPISIRGRTHIHPGVNRSSTSPVSKGPPPLKTPASKSPGEGQTATTSRGAKEPSVK 2280  
Qy 2280 SELSPVARQTSQIGSSKAPSRGSRDSTPSRPAQPLSRPIQSPGRNSISPORNGISPP 2339  
Db 2281 SELSPVARQTSQIGSSKAPSRGSRDSTPSRPAQPLSRPIQSPGRNSISPORNGISPP 2340  
Qy 2340 NKLSQLPRTSPSTASTKSSGSGRMSYTSRQMSQQLTKQTGLSKNASSISPRSESASK 2399  
Db 2341 NKLSQLPRTSPSTASTKSSGSGRMSYTSRQMSQQLTKQTGLSKNASSISPRSESASK 2400  
Qy 2400 GLNOMNGANGANKVELSRMSSTKSSGSGRMSYTSRQMSQQLTKQTGLSKNASSISPRSESASK 2459  
Db 2401 GLNOMNGANGANKVELSRMSSTKSSGSGRMSYTSRQMSQQLTKQTGLSKNASSISPRSESASK 2460  
Qy 2460 ASPESLSPSRPASPTSRQATPVLSPSLPDMSLSTHSSVQAGGWKLPNLSPTIYND 2519  
Db 2461 ASPESLSPSRPASPTSRQATPVLSPSLPDMSLSTHSSVQAGGWKLPNLSPTIYND 2520  
Qy 2520 GRPAKRHDIAHSHSESPSRPINRSGTWKREHSHKSSSLPRVSTWRTGSSSSILSASSE 2579  
Db 2521 GRPAKRHDIAHSHSESPSRPINRSGTWKREHSHKSSSLPRVSTWRTGSSSSILSASSE 2580  
Qy 2580 SSEKAKSEDEKHVNSISGTSQSKENQVSAKGTWRKIKENEFSPNTNSTSQTVSSGATNGAE 2639  
Db 2581 SSEKAKSEDEKHVNSISGTSQSKENQVSAKGTWRKIKENEFSPNTNSTSQTVSSGATNGAE 2640  
Qy 2640 SKTLIYQMAPAVSTEDVWVRIEDCPINNPRSGRSPGTGNTPPVIDSVSEKANPNIKDSKD 2699  
Db 2641 SKTLIYQMAPAVSTEDVWVRIEDCPINNPRSGRSPGTGNTPPVIDSVSEKANPNIKDSKD 2700  
Qy 2700 NQAKQNVGSGVPMRTVGLNRLNSFTQVADPQKGTETKPGQNNPVPVSETNNESSIVER 2759  
Db 2701 NQAKQNVGSGVPMRTVGLNRLNSFTQVADPQKGTETKPGQNNPVPVSETNNESSIVER 2760  
Qy 2760 TPFSSSSSKHSSPSGTVAAARVTPFNPNPSFRKSSADSTARSQIPTPVNNNTKKRDSK 2819  
Db 2761 TPFSSSSSKHSSPSGTVAAARVTPFNPNPSFRKSSADSTARSQIPTPVNNNTKKRDSK 2820  
Qy 2820 TDSTESSGTSQPKRHSGSYLVTSV 2843  
Db 2821 TDSTESSGTSQPKRHSGSYLVTSV 2844

RESULT 8  
US-10-267-502-372  
; Sequence 372, Application US/10267502  
; Publication No. US2004007100A1  
; GENERAL INFORMATION:  
; APPLICANT: Kim, Jaeseob  
; APPLICANT: Galant, Ron  
; TITLE OF INVENTION: Obesity Linked Genes  
; FILE REFERENCE: LSD-07416  
; CURRENT APPLICATION NUMBER: US/10/267, 502  
; CURRENT FILING DATE: 2003-01-27  
; NUMBER OF SEQ ID NOS: 439  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 372  
; LENGTH: 2845  
; TYPE: PR

US-10-267-502-372

Query Match 90.5%; Score 13178; DB 12; Length 2845;  
Best Local Similarity 90.2%; Pred. No. 0;  
Matches 2572; Conservative 111; Mismatches 155; Indels 12; Gaps 10;  
Qy 1 MAAASYDQLLKQVEALKMENSNIROLEEDNSNHLTKLETEASNKKEVLKQOGSTEDSAM 60  
Db 1 MAAASYDQLLKQVEALKMENSNIROLEEDNSNHLTKLETEASNKKEVLKQOGSTEDSAM 60  
Qy 61 AASGQIDLLERLKEINLDSSNPPPGVKLSKMSLSRSGYSGREGSVSRSGECPVPMGSPFR 120  
Db 61 -TSGQIDLLERLKEFNLD-SNPPGVKLSKMSLSRSGYSGREGSVSRSGECPVPMGSPFR 118  
Qy 121 RGFVNGSRESTGYLBELEKERSLLIADLDKBEKEKDWYTAQIQLNLTKRIDSPLTENFSL 180  
Db 119 RTFVNGSRESTGYLBELEKERSLLIADLDKBEKEKDWYTAQIQLNLTKRIDSPLTENFSL 178  
Qy 181 QTDMTROLEYEARQIRVAMEEQLGTCCDMKRAQRRIARIQOIEKDTLRIRQLLQSOAT 240  
Db 179 QTDMTROLEYEARQIRVAMEEQLGTCCDMKRAQRRIARIQOIEKDTLRIRQLLQSOAA 238  
Qy 241 EAERSSQNKHETGSHDAERQNEGQGVGINWATSGNGQSTTRMDHETASVLSSSSTHSA 300  
Db 239 EAERSSQNRHDAASHEAGRCHEGHEGVAESNTAASSSQSPATRVVDHETASVLSSSSTHSA 298  
Qy 301 PRRLTSHLGTKEVWYVYLLSMLGTHDKDDMSRTLLAMSSSDSCISMQSCPLLIQLL 360  
Db 299 PRRLTSHLGTKEVWYVYLLSMLGTHDKDDMSRTLLAMSSSDSCISMQSCPLLIQLL 358  
Qy 361 HGNDKDSVLLGNRSGSKEARARASAAHLNIIHSPQDDKGRREIRVLHLEEQIRAYCETC 420  
Db 359 HGNDKDSVLLGNRSGSKEARARASAAHLNIIHSPQDDKGRREIRVLHLEEQIRAYCETC 418  
Qy 421 WEQBAEHPGDDQDKNMPAPVEHQICPAVCVLMKLSFDEEHRHAMELQGLQAIABLLQ 480  
Db 419 WEQBAEHPGDDQDKNMPAPVEHQICPAVCVLMKLSFDEEHRHAMELQGLQAIABLLQ 478  
Qy 481 VDCMEYGLTNDHYSTILRRYAGMALTNLTFGDVANKATLCMKCMKRALVAQLKSEEDL 540  
Db 479 VDCMEYGLTNDHYSTILRRYAGMALTNLTFGDVANKATLCMKCMKRALVAQLKSEEDL 538  
Qy 541 QOVTASVLRLNSWRADVNSKKTILREVGSVKALMECALVEVKESTLKSVLNMLNSAHT 600  
Db 539 QOVTASVLRLNSWRADVNSKKTILREVGSVKALMECALVEVKESTLKSVLNMLNSAHT 598  
Qy 601 ENKADICAVDGCALAFVGLTYRSTQNTLAIIESGGGILRVNSSLIATNEDHROILRENN 660  
Db 599 ENKADICAVDGCALAFVGLTYRSTQNTLAIIESGGGILRVNSSLIATNEDHROILRENN 658  
Qy 661 CLQTLLOHLKSHSLTIVSNACGTLNLSARNPKQDEALWDMGAVSMKLNLIHSHKHMAM 720  
Db 659 CLQTLLOHLKSHSLTIVSNACGTLNLSARNPKQDEALWDMGAVSMKLNLIHSHKHMAM 718  
Qy 721 GSAALRLNLMANRPAPKVDANIMSPGSSLSPLSHVYRKQKALEAELEDAOHLSETPDINLS 780  
Db 719 GSAALRLNLMANRPAPKVDANIMSPGSSLSPLSHVYRKQKALEAELEDAOHLSETPDINLS 778  
Qy 781 PKASHRSKORHQSILGYDVFDTNRHDDNRSDNFTGNTVLSPLYLNTTVLPSSSSSRGS 840  
Db 779 PKASHRSKORHQSILGYDVFDTNRHDDNRSDNFTGNTVLSPLYLNTTVLPSSSSSRGS 838  
Qy 841 LSSRSSEKORSERERIGLGNVHPATENGTSSKRGQIQTSTAAQIAKYMEEVSAIHTS 900  
Db 839 LSSRSSEKORSERERIGLGNVHPATENGTSSKRGQIQTSTAAQIAKYMEEVSAIHTS 898  
Qy 901 QEDRSSGSTTTELHCVTDERNALRRSSAAHTSNYNTFTKSENSNRCTCMFYAKLEYKRSS 960  
Db 899 QEDRSSGSTTTELHCVTDERNALRRSSAAHTSNYNTFTKSENSNRCTCMFYAKLEYKRSS 958  
Qy 961 NDSLSNVSSSDCYCKRCQMKPSIESYSEDDSEKFCYGOYPADLAHKIHSANEMDDNDGE 1020  
Db 959 NDSLSNVSSSDCYCKRCQMKPSIESYSEDDSEKFCYGOYPADLAHKIHSANEMDDNDGE 1020



2098	QY	NFDWKAQEGANSIVSSLHQ--AAAACILSRQASSDSDSLISLKSIGLSLSPHLLPDQBE	2156
2097	DB	NFDWKAIOEGANSIVSSLHQAAAAACILSRQASSDSLISLKSIGLSLSPHLLTPDQBE	2156
2157	QY	KPFTSNKGRPILKPEKSFLETKKIESESIGIKGGKKVKYSLITGKVRNSNISIGOMKOP	2216
2157	DB	KPFTSNKGRPILKPEKSFLEAKKIESENKIKGGKKVKYSLITGKIRNSNISISOMKOP	2216
2217	QY	LQANMPSISRGTMTHIPVRNNSSSSTSPVSKKGPPPLKTPASKSPBEGOTATITSPGAKP	2276
2217	DB	LPTNMPISISRGTMTHIPGRNNSSSSTSPVSKKGPPPLKTPASKSPBEGCATITSPRGTKP	2276
2277	QY	SVKSELSPVAROTSGTGGSKAPASGSRDSTSPSPAQOPLSRPIOSPGRNISISGRNGI	2336
2277	DB	AGKSELSPIRQTSQISGNGKSSRSGRDSTSPSRPTQPLSRPMQSPGRNISISGRNGI	2336
2337	QY	SPPNKLSQULPRTSSPSTASTKSGSGGMMGYTSPGRQMSOONLTQOTGLSKNASSIIPRSES	2396
2337	DB	SPPNKLSQULPRTSSPSTASTKSGSGGMMGYTSPGRQLSOONLTQOASLKNASSIIPRSES	2396
2397	QY	ASKGLNQMNGANGANKKVELSRMSSTPKSSGESDSRSPVLRQSTFIKEAPSPTLRRKL	2456
2397	DB	ASKGLNQMNGANGSNKKVELSRMSSTPKSSGESDSRSPALVRQSTFIKEAPSPTLRRKL	2456
2457	QY	ESASPELSPSRRPASPTRSOAQTPVLSPSLPDMGLSTHSSVQAGGWKLPNLSPPTIE	2516
2457	DB	ESASPELSPSRRPOSPTPSAQTPVLSPSLPDMGLSTHPSVQAGGWKLPNLSPPTIE	2516
2517	QY	YNDGRPAKHDTARSHSESPSKLPINRSGTWRKHSKHSSSLPRVSTWRTTSSSSSILSA	2576
2517	DB	YNDGRTPKHDTARSHSESPSKLPINRAGTWRKHSKHSSSLPRVSTWRTTSSSSSILSA	2576
2577	QY	SESSSKAKSEDEKHVNSISGTQKQSKENOVASKGTWRKIKENEFFSTNISTQTSVSSGATN	2636
2577	DB	SESSSEKAKSEDERHVSSMPAPQMKENGVPTKGTWRKIKESDISPTGMAQSASGAAS	2636
2637	QY	GAESKTLIIQMAPAVASKTDEVMVRIEDCPINNPRSGRSTPGTNPPIVDSVSEKANPNIXD	2696
2637	DB	GAESKTLIIQMAPPVSKTDEVMVRIEDCPINNPRSGRSTPGTNPPIVDSVSEKSSSIXD	2696
2697	QY	SKDNQ---AKQNVGNGSPMVRTVGLENRNLNSFIQVDAPQOKGTETIKPGQNNPVPVSETNE	2753
2697	DB	SKDSKDTGKQSVGSGS-PVQTVGLETRNLNSFIQVEAPQOKGTETIKPGQSNPVSIAETAE	2755
2754	QY	SSIVERTPTSSSSSKHSSPSGTVVAARVTFFNPNPFRKSSADSTGARPISOIPTPVNNNT	2813
2756	DB	TCIAERTPTSSSSSKHSSPSGTVVAARVTFFNPNPFRKSSADSTGARPISOIPTPVSTNT	2815
2814	QY	KRGDSKTDSTEGSGTQSPKRHSGSYLVTSV	2843
2816	DB	KRGDSKTDITEGSGAQSPKRHSGSYLVTSV	2845

```

US-09-987-482-2
; Sequence 2, Application US/09987482
; Publication NO. US20020184656A1
; GENERAL INFORMATION:
; APPLICANT: BHANDARI, POONAM
; APPLICANT: SHASHIDHARA, L.S.
; TITLE OF INVENTION: IN VIVO ASSAY SYSTEM FOR SCREENING AND VALIDATION OF
; TITLE OF INVENTION: DRUGS AND OTHER SUBSTANCES
; FILE REFERENCE: 056859-0134
; CURRENT APPLICATION NUMBER: US/09/987,482
; CURRENT FILING DATE: 2002-03-21
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 912
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-987-482-2

```

```

Query Match      32.5%; Score 4738; DB 9; Length 912;
Best Local Similarity 100.0%; Pred. No. 3.3e-259;
Matches 912; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 959 SSNDLSNVSSDGYGKRGQKMFPSIESYEDDSKFCSCYQYQYADLAHKSANHHMDND 1018
Db 1 SSNDLSNVSSDGYGKRGQKMFPSIESYEDDSKFCSCYQYQYADLAHKSANHHMDND 60

Qy 1019 GELDTPTNYSGLKYSDEQLNSGROSPONERWAPKHIIEDEIKQSQROSRNQSTTYPY 1078
Db 61 GELDTPTNYSGLKYSDEQLNSGROSPONERWAPKHIIEDEIKQSQROSRNQSTTYPY 120

Qy 1079 TESTDDKHLAFQHPFGQOECVSPYRSGANGSTNRVGNHGINQNVOSLCOEDDYEDD 1138
Db 121 TESTDDKHLAFQHPFGQOECVSPYRSGANGSTNRVGNHGINQNVOSLCOEDDYEDD 180

Qy 1139 KPTNYSERYSEEBQHEEERPTNYSIKYNEBKHHVDPQIDYSLKYATDIPSSQXQSFPS 1198
Db 181 KPTNYSERYSEEBQHEEERPTNYSIKYNEBKHHVDPQIDYSLKYATDIPSSQXQSFPS 240

Qy 1199 KSSSGQSSKTEHMSSESSENTSTPSSNAKRONQLHPSSAQSRGPOKAATCKVSSINQET 1258
Db 241 KSSSGQSSKTEHMSSESSENTSTPSSNAKRONQLHPSSAQSRGPOKAATCKVSSINQET 300

Qy 1259 IQYCYVEDTPICFSRCSLSLSAEDIEGNCQTTQADPANTSANTLOIAEIKKIGTRSAED 1318
Db 301 IQYCYVEDTPICFSRCSLSLSAEDIEGNCQTTQADPANTSANTLOIAEIKKIGTRSAED 360

Qy 1319 PVSEVPVAVSHPRTKSRLOGSSLSSESARHKAVERFSSGAKSPKGAOTPKSPPHYVQ 1378
Db 361 PVSEVPVAVSHPRTKSRLOGSSLSSESARHKAVERFSSGAKSPKGAOTPKSPPHYVQ 420

Qy 1379 ETPLMFSCRTSVSLSDSFESRSIASSVQSPCGMWGSIIPSDLPDPSQGTMPSPRSKT 1438
Db 421 ETPLMFSCRTSVSLSDSFESRSIASSVQSPCGMWGSIIPSDLPDPSQGTMPSPRSKT 480

Qy 1439 PPPPQTAQTKREVKNKAPTAKRESGPKQAANAQVQVLPADTLLHFAFESTPD 1498
Db 481 PPPPQTAQTKREVKNKAPTAKRESGPKQAANAQVQVLPADTLLHFAFESTPD 540

Qy 1499 GFSCSSLSALSDEPFIQKDVLRIMPVPOENDNGNETESEPKNESNQEKEAKTID 1558
Db 541 GFSCSSLSALSDEPFIQKDVLRIMPVPOENDNGNETESEPKNESNQEKEAKTID 600

Qy 1559 SEKDLDSDSDDDIEILEECIIISAMPTKSRKAKPAQATASKLPPVAPKPSQLPVYKLL 1618
Db 601 SEKDLDSDSDDDIEILEECIIISAMPTKSRKAKPAQATASKLPPVAPKPSQLPVYKLL 660

Qy 1619 PSQNRLOPKHVSFTPGDDMPRYVCVEGTPINFSTATSLDLTIESPPNELAAGEVRGG 1678
Db 661 PSQNRLOPKHVSFTPGDDMPRYVCVEGTPINFSTATSLDLTIESPPNELAAGEVRGG 720

Qy 1679 AQSGEPEKRTDIPTEGRSTDEAQQGKTSVTIPELDDNKAEEGDIILAEICINSAMPKGS 1738
Db 721 AQSGEPEKRTDIPTEGRSTDEAQQGKTSVTIPELDDNKAEEGDIILAEICINSAMPKGS 780

Qy 1739 KPRVVKIMDOVOQASASSAPKNOLDGKKKPTSPVKPIPONTYRTRVRKNADSKN 1798
Db 781 KPRVVKIMDOVOQASASSAPKNOLDGKKKPTSPVKPIPONTYRTRVRKNADSKN 840

Qy 1799 LNAERFSDNKSQKQNLKNSKDFNDKLPNNEDVRGSAFDPSPHYTPIEGTPYCFER 1858
Db 841 LNAERFSDNKSQKQNLKNSKDFNDKLPNNEDVRGSAFDPSPHYTPIEGTPYCFER 900

Qy 1859 NDSLSLDFDD 1870
Db 901 NDSLSLDFDD 912

```

```

; GENERAL INFORMATION:
; APPLICANT: BHANDARI, POONAM
; APPLICANT: SHASHIDHARA, L.S.
; TITLE OF INVENTION: IN VIVO ASSAY SYSTEM FOR SCREENING AND VALIDATION OF
; FILE REFERENCE: 056859-0134
; CURRENT APPLICATION NUMBER: US/09/987,482
; CURRENT FILING DATE: 2002-03-21
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 3
; LENGTH: 767
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-987-482-3

Query Match      26.7%; Score 3892; DB 9; Length 767;
Best Local Similarity 100.0%; Pred. No. 1.1e-211;
Matches 767; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAAASYDOLKQVAKWENSLKQELNEDNSNHLTKLETEASNKEVLKQLQGSIEDEAM 60
Db 1 MAAASYDOLKQVAKWENSLKQELNEDNSNHLTKLETEASNKEVLKQLQGSIEDEAM 60

Qy 61 ASSQIDLLERLKEINLDSSNFPQVGLRSKMSLRSYSGREGSVSSRSGECPVPMGSFPR 120
Db 61 ASSQIDLLERLKEINLDSSNFPQVGLRSKMSLRSYSGREGSVSSRSGECPVPMGSFPR 120

Qy 121 RGFVNGRESYGLLEEKERSLLADLDKEEKEKMYAQLQNLTKRIDSLPTENFSL 180
Db 121 RGFVNGRESYGLLEEKERSLLADLDKEEKEKMYAQLQNLTKRIDSLPTENFSL 180

Qy 181 QDTRRQLEYEAPQIRVAMEEQLGTCODMEKRAQRARIQIQIEKDILRIQLQSOAT 240
Db 181 QDTRRQLEYEAPQIRVAMEEQLGTCODMEKRAQRARIQIQIEKDILRIQLQSOAT 240

Qy 241 EAERSSQNKHETGSHDAERQNEGQVGEINWATSGNGQSTTRMDHETASVLSSTHSA 300
Db 241 EAERSSQNKHETGSHDAERQNEGQVGEINWATSGNGQSTTRMDHETASVLSSTHSA 300

Qy 301 PRRLTSHLGTVMVYSLLSMLGTHDKDMSRTLLANSSQSDSCISMRQSGCLPLLIQLL 360
Db 301 PRRLTSHLGTVMVYSLLSMLGTHDKDMSRTLLANSSQSDSCISMRQSGCLPLLIQLL 360

Qy 361 HGNDKDSVLLGNSRSGSEARARASAAALHNTIHSQDDPKRGRREIRVLHLEQIRAYCETC 420
Db 361 HGNDKDSVLLGNSRSGSEARARASAAALHNTIHSQDDPKRGRREIRVLHLEQIRAYCETC 420

Qy 421 WEMOEAHPGMDQKMPAPVEHQICPAVCVLMKLSFDEEHRHAMNELGGLQAIALLQ 480
Db 421 WEMOEAHPGMDQKMPAPVEHQICPAVCVLMKLSFDEEHRHAMNELGGLQAIALLQ 480

Qy 481 VDCMYGLTDHYISITLRRYAGVATNLTFGDVANKATLCMKGMALVAQLKSEEDL 540
Db 481 VDCMYGLTDHYISITLRRYAGVATNLTFGDVANKATLCMKGMALVAQLKSEEDL 540

Qy 541 QCVTASVRLNLSWADVNSKKTIREVGSVKALMECALEVKESTLKSVLGALNLSAHCT 600
Db 541 QCVTASVRLNLSWADVNSKKTIREVGSVKALMECALEVKESTLKSVLGALNLSAHCT 600

Qy 601 ENKADIICAVDGCALAFVGLTLYRSQNTTLAIIESGGGILRNSSLIATNEDHRQILRENN 660
Db 601 ENKADIICAVDGCALAFVGLTLYRSQNTTLAIIESGGGILRNSSLIATNEDHRQILRENN 660

Qy 661 CLQTLQHLKSHSLTIIVSNACGTLWNLSARNPDQOEALWDMGVAMKLNLIHSHKHMIAM 720
Db 661 CLQTLQHLKSHSLTIIVSNACGTLWNLSARNPDQOEALWDMGVAMKLNLIHSHKHMIAM 720

Qy 721 GSAALNLMANRPAPKYKDANIMSPGSSLPSPSHVRKQKALEAELDAQ 767
Db 721 GSAALNLMANRPAPKYKDANIMSPGSSLPSPSHVRKQKALEAELDAQ 767

```



```
QY 1810 DSKQNLKNNKDFNDKPNNEEDVRGSAFADSPHHYPIEGTPYCFSRNDSLSLDFDD 1869
D 1464 DSSQOT-----RTRGDGALQSLCTTPTTEBAYCF-----YDS 1496
QY 1870 DDDVLS-----REKABLRKAKENKSEAKVTSHTLTNSQOQANKTOAIKQPIN---R 1920
D 1497 DEPPATAPPFRASAIPLRAKREKPAGRK-----ETPRAAQPATLPVR 1541
QY 1921 GQPKILQKQSTPQSSKIDPDRGAATDEKLFQNTAENTVCFPSHNSLSLSDIDQENN 1980
D 1542 AQPLI-----VDETPPCYSLTSSASSLSE----- 1566
QY 1981 NKNEPIKETEPDPSQGEPSKQASGVAPKSFHVEDTFCVFSRNSLSLSIDSEDDLIQ 2040
D 1567 -----PEA---PQOPANHARGPQGSQKD-----SSPSPRAEBELIQ 1600
QY 2041 ECISAMP-----KKKPSRLKGDNEKHSRPNMGGITLGEDLTLDKDIORPDSE-H 2090
D 1601 RCISLAMPRTTQVGGRRRKPRALRS-----IRPTEIT-----QKQEEVA 1643
QY 2091 GLSP--DSENFDMKAIQEGANSIVSSILHQAACLSRQASSDSLSILSKGISLGSFP 2148
D 1644 GSDPASLDSVEWQAIQEGANSIVTLHQAARASL--EASSESLSLSLVGSVAGSTL 1701
QY 2149 HLTPDQEKFTSNKGRILKPKGKSTLETKT--EESKIKGKKVKYKSLITGKVR-SN 2206
D 1702 Q-----PSLKRKGRKPAEAGGAWRPEKRTTSTK-----INGSRLFN 1740
QY 2207 SEISQMKQPLQANMPSISGRMTIHPGVNRNSSTSPVSKGKPLKTPASKSPSEGT 2266
D 1741 GPEKAKGTQXMMAGESTMLGRVIV-----SAGPASRTQSKG--ISGPTCTPKTKTGS 1792
QY 2267 ATTPGKAGSVKSELSPVARQTSQIGGSKAKRSRGRSDTSPRPAQPLRPISQGR 2326
D 1793 GTTQPTV-----TKAPSPQQRSSRLHUPCK--SELAALRHPPR 1831
QY 2327 NSISPGRNGISPPNKLQSLPRTSP-----STASTKSG-----SGKMSVTSQROMSQN 2377
D 1832 SATPPARLAKTSPSSSSQTSFASQPLPRRSPPLATPTGGLPGGSLVPKSPARAL---- 1887
QY 2378 LTKQTGLSKNASSI-----PRSESASKGLNQMNNGA--NKKVELSRMS 2420
D 1888 LAKQHKTKSPVRIFFMQRDARRVPPPLARFPEPGRGAGAEPTGARGSLGLVRMA 1947
QY 2421 STKSSGESDESPVLVQSTFKAPSTLARKLEESAPESLSPSRPASPTBSQAQ 2480
D 1948 SARSSGES--SDKSGFRQLTFIKESFG--LLRRRSELSDASTASTQASPRGRPA 2004
QY 2481 TPVL---SPSLPDMSLSTHSSVOAGGWRKLPPLNLSPTIYNDGRPAKRDHIAHSHSESPS 2537
D 2005 LPVFLCSSRCDLURVSPQPLAQRSPQAKPLAPL-----APR-----RTSSESPS 2052
QY 2538 RLPINRGTKWRKSHKSSLSLPRVSTWRRTCSSSSLSASSESEKAKSEDEKXVNSIS- 2596
D 2053 RLPV-RASQRPETVKRYASLPHSIVSRSDSAVSVPPTQANATRGSGDEARPLPRVAP 2111
QY 2597 -GTKQSKENQVSAKGTWRKIKENEFBS-----PTNASTQTSVSGATNGAESKTLIYOMA 2648
D 2112 PGT-----TWRIKQEDVPHILASTLPATLPLRVSSPEDSPAGTPTQ----- 2153
QY 2649 PAVKTEDVWVRIEDCPINPRGRST--GNTPPVIDSVSEKXANPNIKDKONQAKQNV 2706
D 2154 ---RKTSDAVVQTEDVATSKTNSSTSPSLESRDP-----QAP--- 2188
QY 2707 GNGSVPMRTVGLNRLNSFTIQVADQKGTBIRPGQNNPVVSETNESSTIVERTPTSSSS 2766
D 2189 --ASGVAPQSGD-----VDGP-----VLITFPASAPP--HEGLSAVIAGFP----- 2227
QY 2767 SKKHSPPSGTVAARVTPFNPNPSPRKSS--ADSTSPAPSPQIPTPVN 2810
D 2228 TSHGSPSR--AARVPPFNTPVSPMAAATMASDSAVEKAPVSPAS 2271
```

```
RESULT 12
US-10-267-502-371
; Sequence 371, Application US/10267502
; Publication No. US20040071700A1
; GENERAL INFORMATION:
; APPLICANT: Kim, Jaeseob
; TITLE OF INVENTION: Obesity Linked Genes
; FILE REFERENCE: LSD-07416
; CURRENT APPLICATION NUMBER: US/10/267,502
; CURRENT FILING DATE: 2003-01-27
; NUMBER OF SEQ ID NOS: 439
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 371
; LENGTH: 2303
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-267-502-371

Query Match 23.4%; Score 3415.5; DB 12; Length 2303;
Best Local Similarity 34.4%; Pred. No. 8.4e-184; Indels 717; Gaps 94;
Matches 997; Conservative 372; Mismatches 815;

QY 2 AAASYDQLLKQVEALKXVENSRLROLEDNNSHLTKLETEASNMKVLEKQLQGSIEDEA-- 59
D 4 SVAPYEQLVRQVEALKAEHSLRQELADNSHLSKLETSKMEVLKHLQKLEQEARV 63
QY 60 MASSQIDLLRLKELNLDSSNFPVGLKRSKMSLURSTGSRGSGSVSSGECSPVPMGSPF 119
D 64 LVSSQTEVLSQLKALQMDITSLYNLKFPD---TLGPDP---AARTPEGSPV-HGSGP 115
QY 120 RR-GPVNQSRETSGLVLELEKERSILLADLDKEKEKDYVYAOQLNLTKTIDSPLTE-N 177
D 116 SKDSFGLSRATIRLLBELDRERFLNEIEKEKEKLWYSQGLSKLDELPHVETQ 175
QY 178 FSLQDTRRLQLEYEARQIRVAMEEQIGTQODMEKRAQRRIARIQOIEKOILRLRQLQS 237
D 176 FSMQMDLIRQLEFEAQHRSIMEERFGTSDMYQRAIRASLEQIDKELLE----- 228
QY 238 QATEARSSQNHQTHGSHDAERQNEGGVCHINWATSGNGQSTTRMDHTATSVLSSST 297
D 229 -----AQRVQQTPEQALLAV-----KSPVDEPDETEVPHTPE 262
QY 298 KSAPRRLTSLHGTQVMVYSLLSMLGTHDKDDMRITLLAMSSSQDSCISMRQSGCLPLLI 357
D 263 DGTPOGNS---KVEVVFWLLSLMTRDQEDTARTLLAMSSSPESCVMARSGCLPLLI 318
QY 358 QLLHGNKDS---VLLGNSRGSKEARASAAALHNIHSQDDDKRGREIRVHLHLEQIR 414
D 319 QILHTEAAGGRAGAPGAPGAKDARMRAALHNIIVFSQPDQGLARKEKMKVHLVLEQIR 378
QY 415 AYCECTCWEQEAHPGMD-QDKNMPAPVEHQICPAVCVLMKLSFDBEHRHANNELGQL 473
D 379 AYCECTCWMLOARDGGPEGGGAGAPPIEPQICQATCAVMKLSFDEEYRANMELGQL 438
QY 474 ATAEQLQVDCMYGLTNDHYITLRRYAGMALTNLTFGDVANKATLCSMKGCMALVAQL 533
D 439 AVAEQLQVDYEMHKMTRDPLNLRVYAGMPLTNLTFGDVANKATLCARRGCMALVAQL 498
QY 534 KSESDLOQVLTASVLNLSWRADVNSKKTREVGSKVLMCALEVKESTLKVLSALW 593
D 499 ASDSEELQVSSILNLSWRADINSKKVLEASGVTLVQCVRATKSTLKSVALW 558
QY 594 NLSAHTENKADIACVQDGAFLVGLTYRSTQNTLAIIESGGGILNRVNSLIATNEDHR 653
D 559 NLSAHTENKAAICQVQDGAFLVSTLYTKQCSNLAIIESGGGILNRVNSLVRATREDYR 618
QY 654 QILRENCLQTLLOHLKSHSLTIYSNAGTLLWNL-SARNPKDQEAALWDMGVMSLKNLIHS 713
D 619 QVLRDHNCLQTLLOHLTSHSLTIYSNAGTLLWNL-SARSDQELLWDLGAVGMLNVLHS 678
QY 714 KHKMTAMGSAALRNLMANRPAYK--DANIMSPGSSLPFSLHVRKQKALEAELDAQHLSSET 772
```

Db	679	KHKM ANGSAALN LLAHR PAKHQAAAT AVSPGCVPSLVYRQRALEABELDARH LAQA	738
Qy	773	FDNIDN LSPKASHR SKQ-----RHQOSLYGDYVPTD NRHDDN SDN-----PWTG----	817
Db	739	LEHLEKQCP PABAAAT KPLPLRLDGLAQYASDSCG FDDDDAPSLAA AAATGEPAS	798
Qy	818	-----NMTV SPYNT TVLPSSSSRGSLDSRSEKXDRSLERERIGLGN YHPATENPOTS	873
Db	799	PAALS FLGSP LQOQALART PTTRRG--GKEAKD TSGE-----AAVA	840
Qy	874	SKRG QISTTAAQ IAKWEVEYSAI HTSQEDRSSG TTTLHC VTDERNALRBS AAHTHS-	932
Db	841	AKAKA LALAVARID DOLVEDISALHTSD SDSFSUSS-----GDPGQAP REGRAQSCSP	894
Qy	933	NTYNFT SENRTCS NTPYAKLEKRS NDNSLNSVSSDGYGRGQMKP SIEYSYSEDE	991
Db	895	CRGPEGR REAGSRAH PLRLUKAAH ASLNSDLSNGSASDGYC PREHM-----	942
Qy	992	SKFCSYGY PADLAUKI HSANIMODN GELDTP INYSKYDEQ LNSORQSP SONERWAR	1051
Db	943	-----LPCPLA-----ALASR REDP-----	957
Qy	1052	PKHII EIKQEQR QRNOSTTY PVYTES TDDKHLK FQPHGQEQCV SVYRSGANGSE	1111
Db	958	-----RCQPR PSR-----	966
Qy	1112	TNRVSGN GINQVQS LQEDDYEDDK PTNYERYSE EHEBERPTNYS KYNEKR	1171
Db	967	-----LDL LPGCAE PPAREAT SADARVAT-----IKLSPT YQ	1000
Qy	1172	HVDQ PIDYSLK YATDIP SSOKOSF SFGSSSQSSK TEHMS SSSENT TPSSNAKRON QOL	1231
Db	1001	HV-----PLLEG SRA GAEP LAGFI-----SFGARK QAWL	1031
Qy	1232	HPSSAQ SGOPQK AATCKVSS INQETI QTYC VEDT PFCR CSSLSL SSLA E BIGCQ	1291
Db	1032	---PADHLS KVPEK LAAAPL-SVASKALQ KLAQEG PLSRCS SSLS SSSA-GRPG PSE	1086
Qy	1292	TTQ EADSANT QIAEIK EKIGTR SAB-----DPVSE VPASVQ HPRTKSR LQSS LSS	1344
Db	1087	GGDL DSD SSSL-----EGLE EAGP SEALD STWRAP GATSLP VAIPAPR-----RNRG GLGV	1139
Qy	1345	EGARH KAVEF SSGAKSP SGAQTP KSPPEH VQETPL MF SRCT SVSS LDG FEGRS IASS	1404
Db	1140	EDA-----TSSSEN VQETPL VLRCS SVSS SLG SFESP SIASS	1179
Qy	1405	VOSEPC SGMVSGI ISPSDL PDSPG QTMPPSR SKTTP-----PPPG TATKREV PNKAPTAEK	1462
Db	1180	IPSEPC SGQSGT ISPS ELPDSPG QTMPPSR SKTTP LAPAQG-----PPEATQ FSLQ	1232
Qy	1463	RESGPKQAAVNA VQVQLP-----DADTL HFEAT STPD GFCSS SLGALS LDEPFI OKD	1519
Db	1233	WESYV KRF LDIAD CHRC KLPS ELDAGS--VRTV TEKPDEN FSCASS LSALAH EHYVQD	1291
Qy	1520	VELRIM PPYQ ENDNGNET EQPKES NENQ EAEK TIDSEK DLDD DDDD DIEIL ECI	1579
Db	1292	VELRL LSPAC PERGGAGG AGLHF AGHRR REEGP APTGS RRP-----GAADQ LELLRECL	1347
Qy	1580	ISAMPT KSRKAK PAQTASK LPPV VAPKPSQ LPVYK LLPSON RLQPKHVS FTGDDMP	1639
Db	1348	GAAPV PARL K-----VASAL VP-----GRRAL PVPV TMLV FAPAPAQ-----EDDS	1388
Qy	1640	RVYCV EGT PINF STAT SLD LIES PPNELA AGEV RGGAQ SGE PEK DTP TEGR STDE	1699
Db	1389	CYD S AEGT PVNF SSAAS LSD ELTQ PPRDPQ PGPAGR-----QRP T-----GRPT SA	1435
Qy	1700	AQGGT SSVTI PELDDN KAE EGD ILAE CINS AMPK GKSHK PFRV KKIMD VQVQ QAS SSA	1759
Db	1436	ROA-----MGRH K-----	1451
Qy	1760	PNKNQ LDGK K-----KKPT SPVK PIPQ NTEY TRVR KNAD SKN LN AER VPS DNKD	1810
Db	1452	AEQSRGAGK NRAGLE LPLGR PPS GAPA-----	1480
Qy	1811	SKQNLK NNSK DFNDK LPNNED VRGVS FADPS PHV TPT EGT YCF SRNDS LSLDF DDD	1870
Db	1481	GSFPG-----RTRGDGALQSLCLT TPT EEAVY CFY GNS-----	1514
Qy	1871	DVDSLREK EALRKAK NKNS EAKV TSH TELTS NQO SANK TOA AKQ P NRG QPK LOKQ	1930
Db	1515	-----D BEP PAAAPT TH-----RRTS ALPR-AFT REPR-----QGR	1545
Qy	1931	STPQSSK DIPRGAAT DEKLO-NF A ENTPV CF SHNS LSLSD DIQ ENN K EN EPIKE	1989
Db	1546	KEAPAS KAP--A P PARTQ SLAD ETPPCY SLSS SASSLS-----E P P-----	1591
Qy	1990	TEP PD SQE PSK PQAS GYAPK SFH VEDT PVC FERN SSLS SLSD SD EDD LLQ ECI SSAMPK	2049
Db	1592	SEPPAVH PRGRE PAVT-----KDPG GGGRD SSPSP--RAA ELLQ RCI SSALPR	1639
Qy	2050	KKXP-S ELK DNK EKH SPRNMG GILG EDLT L DLK DIQ RP--DSEHG-----LS PDSEN	2098
Db	1640	RRPVS GLR-----RRKPR-----ATRLD-----ERP AGSR ER GEE BAAGS DRAS DLDS	1683
Qy	2099	FDWKATQ EGANS IVSS LHQAAA ACLSRQAS SDSD SILSK GIS LGS-----P H LTPDQ	2154
Db	1684	VENRAI QEGANS IVTWL HQAAA--TREAS SED SILSV FVSG SLVSGT LQPK HKG GRQ	1740
Qy	2155	BEK PFT NKG PRILK PGEK STLET TKIE SEK IGK GK VYK SLIT TKVRSN SEI SQMK	2214
Db	1741	AEEMGS ARRP-----EKGAAS VKTSG SPRPA GPEK-----PRGT	1777
Qy	2215	QPLQAN MPST SRGT MIHI-----PGRVNS SSSTSPVSK KGP LKP TPA--SKS	2260
Db	1778	QKTTG VPV LVRG TVIYV PSPAPAPQK TGP PRAT PKVAP-----PCLAQ PAPAKV	1832
Qy	2261	PSE QTATTS--PRGAKP SVKSELSPVARQTS QIGSSK APSRG SGR DST PSRPAQ DPLSR	2319
Db	1833	PSPQQR SRSLHR PAKTS ELATLS QPPRSATP PARLAKT PSSSS QTS-----PASO PLPR	1888
Qy	2320	---PIQSP GRNS ISPC ENGIS PPNKL SQLPRT SSPSTASTK SSGSGK MSYTS PGRQMS QO	2376
Db	1889	KRPV TQAA-----GALP GPGAS VPV EKT FARTILAK QHK-----TORS PVR PFMW	1934
Qy	2377	NLTKQT GLSKN VASSI P SASAK GLQNM NNGA--NKKV ELSRMS TKSSG SSDSR SERP	2435
Db	1935	RPARR-GPP LARAVP--EPGPR GRAGTEAGP GARGGRLGLVRV ASALSSG SES--SDRS	1989
Qy	2436	VLRQSTFI KEASPT LARKL EBSAF ESLS PSR PAS PTRSQAQ TPVLS PDLK SLST	2495
Db	1990	GFRQLT FIKE--SPGLRRR SELSSA ESAASAPQGA SPRGR-----PALP AVFLCS	2040
Qy	2496	HSSVQ-----AGGWK LPPNLS PTI EYND CRP AKR HDIAR SHSES PSRL PINRSGTW	2547
Db	2041	SRCE LPAAPRQ GAPAPARQ PPAARPS--PGERPAR-----RTTSE SPSEL PV-RAPAA	2091
Qy	2548	KREH SKHSS SLPRV TWR TGSSSS L SAS SESEKAK SEDEK HVN SISG TKSKENQ VS	2607
Db	2092	RPETV KYASLPHI SVARRPD GAVPAPAPAGQ LGLSDVDG PGLAKAPI--SAPFV HEGLG	2140
Qy	2608	AKG-TW EKI ENEF-----SPTN STQTVSSG ATNGAES KTLIVQMA PAVSK TEDV WVRI	2661
Db	2141	APGTTW RIR DEDDVPH LRSTL FATALP RGS TPE DAPA-----GPPR KTS DAVQOT	2193
Qy	2662	EDCP INNPR SGR SPT GMT--PPVIDS VSEK ANPNIKD SKDNQAKQNV GNSVPMR TVGLE	2719
Db	2194	EVAAPK TNS TSPS LETRE FPGAPAGQ LGLSDVDG PGLAKAPI--SAPFV HEGLG	2250
Qy	2720	NRLNS FQVDA PQKGT EIKP GQNNP VPV SETH ESSIV ERTP FSSSS SKHSS P SGTVA	2779
Db	2251	VAVGGF-----PASR HGSPSR--SA	2268
Qy	2780	RVTPFN NPN SPR KSSAD STGA	2800
Db	2269	RVPEFN VPSPMVVAAT TDSA	2289

RESULT 13  
US-10-267-502-368  
; Sequence 368, Application US/10267502  
; Publication No. US20040071700A1  
; GENERAL INFORMATION:  
; APPLICANT: Kim, Jaeseob  
; TITLE OF INVENTION: Obesity Linked Genes  
; FILE REFERENCE: LSD-07416  
; CURRENT APPLICATION NUMBER: US/10/267,502  
; CURRENT FILING DATE: 2003-01-27  
; NUMBER OF SEQ ID NOS: 439  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 368  
; LENGTH: 2417  
; TYPE: PRT  
; ORGANISM: Drosophila melanogaster  
US-10-267-502-368  
Query Match 12.1%; Score 1765; DB 12; Length 2417;  
Best Local Similarity 25.3%; Pred. No. 2.2e-90;  
Matches 719; Conservative 325; Mismatches 838; Indels 964; Gaps 100;  
QY 243 ERSQKHETGSHDAERON-EGQGVGEINMATSGGGSTTR-----MDHETASVLSSS 295  
DB 129 ELREMRHRLDRNFERQSAQQQDLDELPPRNGGSPASAGRPSRSGKPSYTLRSLFDGD 188  
QY 296 STHSAPR-----RLTSH-----LGTQVEMVYSLLSMLGTHDKDDMSRLLAM 337  
DB 189 APAPAPLPKGAWTTSDFDERYTSSAVATLGSKEVCVSLLSMLGSDPLEWAKKFLLEL 248  
QY 338 SSGODSCISMRQGGCLPLLQLLHGNDKDSVLLGNRSGSKAPARASAAALHNIHSQPD 397  
DB 249 SGAQSCATLRRSGCMPLLVQWHPDND-----QEVKCKAGQALHNVVHSHPD 298  
QY 398 KRGRRETRVHLEOTRAYCETCWENQEAHEPGM--DQDKNMPAPVHEQICPACVLMK 455  
DB 299 KAGREAKVRLDQIVDYCSFLKTLQSGGEIADDSRHPL-----AATSSLMK 349  
QY 456 LSFDEEHRHAMNELGGLOAIALLQVDCENYMG-LTNDHYSITLRRYAGMALTNLTGDDVA 514  
DB 350 VSFDEEHRHAMCELGAHAIPLNVLHDHVAHVGKPEDQCCNLSRLRYALMALTNLTGDD 409  
QY 515 NKATLCSMKCMALVAQLKSESDILOQVATSVLRNLSHRADVNSKTLREVGSVKALME 574  
DB 410 NKALLOQKQFMALVAQLDSAPDDLQVTVASVLRNLSWRADSNMKAVLNEIGTVTALAL 469  
QY 575 CALEVKKESTLKSVLNALMSLHACTENKADICAVDGAFLAVGTITYSQNTNLAIIES 634  
DB 470 AAMNRSNTLKALLSALNLSAHCTNKAEFCVAVDGAFLAVGMLSYEGSPKTLKIEN 529  
QY 635 GGGILRVSSLIATNEDHRLKRENCLOTLQHLKSHSLITIVSNACGTLWNLARNPKD 694  
DB 530 AGGILRVSSHIAVCEPYQILRHQNCALTLQQLKSESLITVVSNSCGTLWNLARSAD 589  
QY 695 QEALWDMGAVSMLKNLHSHKMIAMGSAALNLMANRPAYK-----DANIMSPG-SSL 749  
DB 590 QKFLWDNGAVPMRLSLHSHKAMISGSSALKNLNFRAVQNHQLDPIARSMGLKAL 649  
QY 750 PSJLVKQKALEAELDAQHLSETFDNIDNLSPKASHRSKQRHKQSLYGDYVDTNRHDDN 809  
DB 650 PTLKARKAKALQQLGERHTAETCDNL----- 677  
QY 810 RSDNFNTGNVTLSPYNTTTLVSSSSSRGSLDSRSKXDRSLERERGLGNHYHPATEN 869  
DB 678 -----TGG-----KLDKFRASSRRHP----- 695  
QY 870 PGTSSKRGLOITTAQIAQKAMEVSAIHTSQDRSSGSGTTELHCVTDBERNALRRSSAAH 929  
DB 696 -----APRLTR-----SAMLTKSRSRDSVTSKSDCAYDH----- 725

QY 930 THSNTYNTFTSGNSRNTCTMPYAKLEVYKRSNDLSNVSSSDGYGKRGQMKPSTESYSED 989  
DB 726 -----LIRSASASDAHRK-----VKPKITDLDLE 749  
QY 990 DESKFCYSQYPADLAHKIHSANHMNDNDGELDTPTINYSLKYSDLEQNSGRQSPQNERW 1049  
DB 750 ME-----QDTEATEEQPIDYSVKYS----- 769  
QY 1050 ASPKHIIIEDEIKQSFQRORNCQSTTYPVYTESTDDKHLKTFQHFQEQECVSPYRSRGANG 1109  
DB 770 -----ENATKISTY----- 778  
QY 1110 SETNRVGSNHGINQVNSQSQEDDYDDKPTNYSERYSEEEHQBEEBEPRTNYSIKYNEE 1169  
DB 779 -----QETDL--DQPTDFSLYAE-----NQIESDLDISGPAGGQ 811  
QY 1170 KHVDDQPIDYSLKYATDIP--SSQKQSFSSKSSSQSQSKTEH-----MSSSENSTSTPSS 1223  
DB 812 KSTITPP-----AETVPEKSEGEILLILDDSVKCYQTEDTPYVISNAASVTDLRVAA 864  
QY 1224 NAKRQNLHPS--SAQSRSGQFQ--KAATCKVASINQETIQTICYVEDTICFSCSSLS 1278  
DB 865 KADAEAEVKPEVREVTSGEGAPKLPKLSQCGSGSYTPKEPYNICEBGTGPGYFSKYDLS 924  
QY 1279 SLSAEDBIG--CNQTTQEDASANTLQIAEIKKIGTRSAEDPVSEVPAVSOHPKTSKRL 1337  
DB 925 SL-----DESGKANAIVGTD-----ADIKPLEKQEBEQSPAEQVLTAKPPTQANSAL 973  
QY 1338 QGSSLSSESARHKAVERFSGAKSPSKGAQTPKPPPHYQETPLMFSRCTVSSLSDFE 1397  
DB 974 -----ETPLMFSRSMDSLVHDP 992  
QY 1398 SRSTA-----SSVQSEPCSGMVSGIISPSDLPSPGQTMPPSRSKTTPPPQTAQTKREV 1452  
DB 993 DVDVANCDDKSSVVD--FSRLASGVISSFIPDPSFTQSPMQS----- 1033  
QY 1453 PKNKAPTAERESGPKQAAVNAQVQVLPDADTLHFAETESTPDGFCSSLSALS-- 1511  
DB 1034 PRNSVAGSGQVDSPPVPIPASLOPLRSVPE--DDLSSFVVEHTPAQFSTATSLNSLIV 1092  
QY 1512 -DEPFIQKDELRIIMPVQENDNGNE-----TESEQPKSENQEK--AEKT 1556  
DB 1093 DDE-----KAPASVAEDNEDELLANCINMGQKPTKAVKSTVNVSEVDVAET 1143  
QY 1557 IDS-----EKD-----LLDD--SDDDDI-- 1572  
DB 1144 IRSYCTEDTPALLSKVPSNTNLVSISSSTDPDKATAGQAQMAHQSDDVSSNASCGG 1203  
QY 1573 ---EILEECIISAMPTKSSRAKKAQTAGKLPVPAVK--PSQLPVYKLLPSONRLQPK 1628  
DB 1204 ASGHLQCCIRDM-----KKPLGEATSDPTAMLRGCGNELPGY--LPS----- 1245  
QY 1629 HVSTPGDDMPRVYCVGEPTINPSTATSLDLTIES-----PPNELAAGEVGRGAQSGEF 1684  
DB 1246 -----ADEMNK--FLVEDSPCNFVSWGLSNTLVGSSLVGFAVOLKETE--PSSADQNP 1297  
QY 1685 EKRTIPTEGSTDEAQQGKTSSVTIPELDDNKAEBGDILAEC----- 1727  
DB 1298 KAKPGKQOVRPHQWDDSLSSLSIDSEDDTNLLSQAIANGCNRPKSNLGFSSNGKSS 1357  
QY 1728 -----INSAMPKSKHKPFRVKIMQVQAQASASSAPKNQ-----LDG-- 1767  
DB 1358 SLSQSQPIAINAATASGLSNTAMTVRSQQQESYSSVSDSDNDNQSKSLFELCILKGMV 1417  
QY 1768 KKKXP-----TSPVKPIQNTTEYRT-----RVRKNADSKNLNA 1801  
DB 1418 KTKPGARAQVQVQPIVGSVSSVSNFSLQKQFDSLVLQVLPSSGQVKQORHHHHHRRER 1477  
QY 1802 ERVFSNDKDSK--KONLKN--SKDFNDKLPNNEDRVGSPAFSPHHYTFIEGTPYCFSR 1858  
DB 1478 ER--ERRKDEKLLQECINTGISKKIN-AVPKNV--LATSAAALEPCH--PMAAT----- 1523  
QY 1859 NDSLSSLDFFDDDDVDLREKAEALPKAKENKSEAKVTSHTELTSNQOSANKTQAIKQPI 1918







QY	2702	AKQNVGNGSVPMRTVTVGLENLNSFIQVDAPDQKGTETKPGQNNPVPVSETNESSIVERTP	2761
Db	1566	AKAPI---SAPFVHEGLGVAVGGF-----	1586
QY	2762	FSSSSSKHSSPSGCTVAAAVTFFNNYNSPRKSSADSTSA	2800
Db	1587	-----PASRHGSPSR--SARVPPFNYPSPMVAATTDSA	1619

Search completed: August 25, 2004, 17:53:02  
Job time : 215.5 secs

***This Page Blank (uspto)***

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 25, 2004, 17:08:42 ; Search time 37.5 seconds  
(without alignments)  
3913.937 Million cell updates/sec

Title: US-09-442-489F-7  
Perfect score: 14566  
Sequence: 1 MAAASYDQLLKQVEALKMEN.....ESSGTQSPKRGSGSYLVTSV 2843

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents\_AA\*  
1: /cgn2\_6/ptodata/2/iaa/5A\_COMB.pep:\*  
2: /cgn2\_6/ptodata/2/iaa/5B\_COMB.pep:\*  
3: /cgn2\_6/ptodata/2/iaa/6A\_COMB.pep:\*  
4: /cgn2\_6/ptodata/2/iaa/6B\_COMB.pep:\*  
5: /cgn2\_6/ptodata/2/iaa/PCTUS\_COMB.pep:\*  
6: /cgn2\_6/ptodata/2/iaa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	14566	100.0	2843	1	US-08-452-655B-2
2	14566	100.0	2843	1	US-08-452-655B-7
3	14566	100.0	2843	3	US-08-450-582-2
4	14566	100.0	2843	3	US-08-450-582-7
5	14566	100.0	2843	3	US-08-821-355A-7
6	14566	100.0	2843	2	US-09-003-687A-7
7	14566	100.0	2843	3	US-09-136-605-7
8	14548.5	99.9	2842	1	US-07-741-940-7
9	14548.5	99.9	2842	1	US-08-289-548A-7
10	14548.5	99.9	2842	4	US-08-452-654-7
11	14548.5	99.9	2842	4	US-08-449-731-7
12	14533	99.8	2843	1	US-07-741-940-2
13	14533	99.8	2843	1	US-08-289-548A-2
14	14533	99.8	2843	1	US-08-452-654-2
15	14533	99.8	2843	2	US-08-370-235A-2
16	14533	99.8	2843	4	US-08-449-731-2
17	507.5	3.5	178	2	US-08-630-822A-100
18	507.5	3.5	178	2	US-09-005-069-100
19	507.5	3.5	178	4	US-09-171-156A-49
20	507.5	3.5	178	4	US-09-004-730A-49
21	507.5	3.5	178	4	US-08-981-799A-49
22	502.5	3.4	2137	4	US-09-134-001C-4463
23	461.5	3.2	10182	4	US-09-134-001C-3159
24	413	2.8	3696	4	US-09-134-001C-5080
25	385	2.6	8991	4	US-08-714-741-32
26	382.5	2.6	3256	4	US-09-519-172-98
27	382.5	2.6	3256	4	US-09-976-594-22

## RESULT 1

US-08-452-655B-2  
; Sequence 2, Application US/08452655B  
; Patent No. 5783666

### GENERAL INFORMATION:

APPLICANT: ALBERTSEN, HANS  
APPLICANT: ANAND, RAKESH  
APPLICANT: CARLSON, MARY  
APPLICANT: GRODEN, JOANNA  
APPLICANT: HEDGE, PHILIP J.  
APPLICANT: JOSLYN, GEOFF  
APPLICANT: KINZLER, KENNETH  
APPLICANT: MARKHAM, ALEXANDER F.  
APPLICANT: NAKAMURA, YUSUKE  
APPLICANT: THLIVERIS, ANDREW  
TITLE OF INVENTION: INHERITED AND SOMATIC MUTATIONS OF APC  
TITLE OF INVENTION: GENE IN COLORECTAL CANCER IN HUMANS  
NUMBER OF SEQUENCES: 102  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Banner & Witcoff, Ltd.  
STREET: 1001 G Street, NW  
CITY: Washington  
STATE: D.C.  
ZIP: 20001-4598  
COUNTRY: USA  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/452,655B  
FILING DATE: 25-MAY-1995  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/289,548  
FILING DATE: 12-AUG-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/741,940  
FILING DATE: 08-AUG-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Kagan, Sarah A.  
REGISTRATION NUMBER: 32,141  
REFERENCE/DOCKET NUMBER: 1107.49964  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-508-9100  
TELEFAX: 202-508-9299  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2843 amino acids

## ALIGNMENTS

28	380.5	2.6	3248	1	US-08-353-700-1	Sequence 1, Appli
29	380.5	2.6	3248	5	PCT-US95-16216-1	Sequence 1, Appli
30	377.5	2.6	1581	3	US-09-110-517-2	Sequence 2, Appli
31	370.5	2.5	3969	3	US-08-061-376-5	Sequence 5, Appli
32	363	2.5	2468	4	US-09-976-594-726	Sequence 726, App
33	361	2.5	2662	4	US-09-595-684B-31	Sequence 31, Appli
34	346	2.4	2954	4	US-09-150-867-1	Sequence 1, Appli
35	337.5	2.3	2185	4	US-09-854-856-36	Sequence 36, Appli
36	337.5	2.3	2245	4	US-09-854-856-4	Sequence 4, Appli
37	337.5	2.3	2322	4	US-09-854-856-34	Sequence 34, Appli
38	337.5	2.3	2382	4	US-09-854-856-2	Sequence 2, Appli
39	332.5	2.3	2482	1	US-08-328-254-6	Sequence 6, Appli
40	322	2.2	2157	4	US-09-854-856-52	Sequence 52, Appli
41	322	2.2	2217	4	US-09-854-856-20	Sequence 20, Appli
42	322	2.2	2294	4	US-09-854-856-50	Sequence 50, Appli
43	322	2.2	2354	4	US-09-854-856-18	Sequence 18, Appli
44	313.5	2.2	1596	4	US-08-978-277A-4	Sequence 4, Appli
45	308	2.1	1969	4	US-09-418-710-72	Sequence 72, Appli

TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-452-655B-2

Query Match 100.0%; Score 14566; DB 1; Length 2843;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2843; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	MAAASVDQLLKQVEALQWENLNRLQELDNHNLTKLETEASNMKEVLKQLOGSIDEAM	60
Db	1	MAAASVDQLLKQVEALQWENLNRLQELDNHNLTKLETEASNMKEVLKQLOGSIDEAM	60
Qy	61	ASSGQIDLLRLKELNLDSSNFFGVKLRKMSLRYSRGSGVSSRGSCSPVPMGSPFR	120
Db	61	ASSGQIDLLRLKELNLDSSNFFGVKLRKMSLRYSRGSGVSSRGSCSPVPMGSPFR	120
Qy	121	RGFVNGSRESTGYLEELKERSILLADLDKEEKEDWYQAQLNLTNRIDSLPLTENFSL	180
Db	121	RGFVNGSRESTGYLEELKERSILLADLDKEEKEDWYQAQLNLTNRIDSLPLTENFSL	180
Qy	181	QDTMTRRQLEVEARQIRVAMEEQLGTCQDMKEKAQRRIARIQIQIEKDILIRIQLLOSQAT	240
Db	181	QDTMTRRQLEVEARQIRVAMEEQLGTCQDMKEKAQRRIARIQIQIEKDILIRIQLLOSQAT	240
Qy	241	EAERSSQNHKETSCHDAERQNEGGVGEINMATSNGQSGSTTRMDHETASVSSSSSTHSA	300
Db	241	EAERSSQNHKETSCHDAERQNEGGVGEINMATSNGQSGSTTRMDHETASVSSSSSTHSA	300
Qy	301	PRRLTSLHGTKEVMYVSLLSMLGTHDKDMSRTLLAMSSQSCISMRQSGCLPLLIQLL	360
Db	301	PRRLTSLHGTKEVMYVSLLSMLGTHDKDMSRTLLAMSSQSCISMRQSGCLPLLIQLL	360
Qy	361	HGNDKQSVILGNRSGSKEARASAAALHNIHISQPDQKGRREIRVLHLEQIRAYCETC	420
Db	361	HGNDKQSVILGNRSGSKEARASAAALHNIHISQPDQKGRREIRVLHLEQIRAYCETC	420
Qy	421	WEWQEAHEPGMDQKNPMPAPVHQICPAVCVLMKLSFDEEHRHAMNELGGQAIAELQ	480
Db	421	WEWQEAHEPGMDQKNPMPAPVHQICPAVCVLMKLSFDEEHRHAMNELGGQAIAELQ	480
Qy	481	VDCEMYGLTNDHYSITLRRYAGMALNTLFGDVANRATILCSMKGMCRALVAQKSEDL	540
Db	481	VDCEMYGLTNDHYSITLRRYAGMALNTLFGDVANRATILCSMKGMCRALVAQKSEDL	540
Qy	541	QOVVASVLRNLWRADVNSKKTIREVGSVKALMECALEVKKESTLKSVALWNLSAHCT	600
Db	541	QOVVASVLRNLWRADVNSKKTIREVGSVKALMECALEVKKESTLKSVALWNLSAHCT	600
Qy	601	ENKADICAVDGAALAFVLGTLTYRSQNTLAIIESGGGILNRVSSSLIATNEDHRQILRENN	660
Db	601	ENKADICAVDGAALAFVLGTLTYRSQNTLAIIESGGGILNRVSSSLIATNEDHRQILRENN	660
Qy	661	CLOTLLQHLKSHSLITIVSNACGLTNLWLSARNPKDQALWDMGAVSMKMLIHSKHNIAM	720
Db	661	CLOTLLQHLKSHSLITIVSNACGLTNLWLSARNPKDQALWDMGAVSMKMLIHSKHNIAM	720
Qy	721	GSAALRNLMANRPKAKDANIMSGSSLPVSLHVRKOKALELDAHLSETFDNIDNLS	780
Db	721	GSAALRNLMANRPKAKDANIMSGSSLPVSLHVRKOKALELDAHLSETFDNIDNLS	780
Qy	781	PKASHRSKQHKQSLYGDYVFTNRHDDNPSDNFTGNMTVLSPLYNTTVLSPSSSSRGS	840
Db	781	PKASHRSKQHKQSLYGDYVFTNRHDDNPSDNFTGNMTVLSPLYNTTVLSPSSSSRGS	840
Qy	841	LDSSRSKDSLRERERGIGLVNHPATENPGTSSKRGLOISTTAAQIAKWEVEVAIHTS	900
Db	841	LDSSRSKDSLRERERGIGLVNHPATENPGTSSKRGLOISTTAAQIAKWEVEVAIHTS	900
Qy	901	QEDRSSGSTTELHCVTDERNALRRSAAHTHSNTYNTFTKSENRTCSMPYAKLEYKRSS	960
Db	901	QEDRSSGSTTELHCVTDERNALRRSAAHTHSNTYNTFTKSENRTCSMPYAKLEYKRSS	960

Qy	961	NDSLNSVSSSDGYGKRGQMKFPIESYSEDDDESKFCSYGOYPADLAHKIHSANHMDNDGE	1020
Db	961	NDSLNSVSSSDGYGKRGQMKFPIESYSEDDDESKFCSYGOYPADLAHKIHSANHMDNDGE	1020
Qy	1021	LDTPINSLKYSDEQLANSRGSPQNERWARPPKHIIIEDEIKQSEQRQSRQSTTYPVYTE	1080
Db	1021	LDTPINSLKYSDEQLANSRGSPQNERWARPPKHIIIEDEIKQSEQRQSRQSTTYPVYTE	1080
Qy	1081	STDDKHLKFKQHFQGCQECVSPYRSGANGSETNRVGNHGINQNVQSLOQEDDYEDDKP	1140
Db	1081	STDDKHLKFKQHFQGCQECVSPYRSGANGSETNRVGNHGINQNVQSLOQEDDYEDDKP	1140
Qy	1141	TNYSERISDEEQQHEEERPTNYISIKYNEEKHVDQPIIDYSLKYATDIPSSQKQSFSSKS	1200
Db	1141	TNYSERISDEEQQHEEERPTNYISIKYNEEKHVDQPIIDYSLKYATDIPSSQKQSFSSKS	1200
Qy	1201	SSGSSKTEHMSSESSENTSTPSSNAKQNLHPSSAQSRGQPOKAATCKVSSINQETIQ	1260
Db	1201	SSGSSKTEHMSSESSENTSTPSSNAKQNLHPSSAQSRGQPOKAATCKVSSINQETIQ	1260
Qy	1261	TYCVEDTPICFSRCSLSISSAEDTGCNQTTQEDSANTLQIAETKEKIGTRSAEDPV	1320
Db	1261	TYCVEDTPICFSRCSLSISSAEDTGCNQTTQEDSANTLQIAETKEKIGTRSAEDPV	1320
Qy	1321	SEVPAVSOHPPTKSRRLQSGSLSSSESARHKAVERPSSGAKSPSKSGAOTPKSPPEHYVOET	1380
Db	1321	SEVPAVSOHPPTKSRRLQSGSLSSSESARHKAVERPSSGAKSPSKSGAOTPKSPPEHYVOET	1380
Qy	1381	PLMFSRCTSVSSLDSPESRSIASSVQSEPCSGMVSGIISPSDLPDSQGTMPSPRSKTPP	1440
Db	1381	PLMFSRCTSVSSLDSPESRSIASSVQSEPCSGMVSGIISPSDLPDSQGTMPSPRSKTPP	1440
Qy	1441	PPPTAQTKREVPKNKAFTAKESGPKQAANNAVQVLPDADTLLHFAETESTPDGF	1500
Db	1441	PPPTAQTKREVPKNKAFTAKESGPKQAANNAVQVLPDADTLLHFAETESTPDGF	1500
Qy	1501	SCSSLSALSDELPEFIQKDVLELRIMPPVQENDNGNETESQPKESNENQEKAKTIDSE	1560
Db	1501	SCSSLSALSDELPEFIQKDVLELRIMPPVQENDNGNETESQPKESNENQEKAKTIDSE	1560
Qy	1561	KDLDDDDDDDDIEILEECIIISAMPTKSSRKAKPAQATASKLPPPVARKPSQLPVYKLLPS	1620
Db	1561	KDLDDDDDDDDIEILEECIIISAMPTKSSRKAKPAQATASKLPPPVARKPSQLPVYKLLPS	1620
Qy	1621	QNRLOPKQHSYFTPGDDMPRVYCVGPTINFSTATSLDITIESSPNELAAAGVGRGAQ	1680
Db	1621	QNRLOPKQHSYFTPGDDMPRVYCVGPTINFSTATSLDITIESSPNELAAAGVGRGAQ	1680
Qy	1681	SGFEFEKRDITPTGRTSTDEAQQGKTSVTIPELDDNKAEBGDIILAEICINSAMPKGSHP	1740
Db	1681	SGFEFEKRDITPTGRTSTDEAQQGKTSVTIPELDDNKAEBGDIILAEICINSAMPKGSHP	1740
Qy	1741	FRVKIKIDVOQOASASSAPNKLQDCKKKKPTSPVKPIPQNTYRTRVRKNADSKNNLN	1800
Db	1741	FRVKIKIDVOQOASASSAPNKLQDCKKKKPTSPVKPIPQNTYRTRVRKNADSKNNLN	1800
Qy	1801	AERVFSDKDSKKQNLKNNKDFNDKLPNNEDRVGRSFAFDSPHHYTPIEGTYPCFSRND	1860
Db	1801	AERVFSDKDSKKQNLKNNKDFNDKLPNNEDRVGRSFAFDSPHHYTPIEGTYPCFSRND	1860
Qy	1861	SLSLDDDDDDVLSREKAELEKAKENKSEAKVTSHTLTSSQOSANKTQAKOPINR	1920
Db	1861	SLSLDDDDDDVLSREKAELEKAKENKSEAKVTSHTLTSSQOSANKTQAKOPINR	1920
Qy	1921	QCPKPIILQKQSTFPQSSKDIIPDRGAATDEKILQNPFAIENTPVCFSHNSLSLSDIOENN	1980
Db	1921	QCPKPIILQKQSTFPQSSKDIIPDRGAATDEKILQNPFAIENTPVCFSHNSLSLSDIOENN	1980
Qy	1981	KNENEPKETEPPDPSQESPKQASGVAPKSFHVEDTPVCFSRNSLSLSIISSEDDLQ	2040
Db	1981	KNENEPKETEPPDPSQESPKQASGVAPKSFHVEDTPVCFSRNSLSLSIISSEDDLQ	2040
Qy	2041	ECISSAMPKKKKPKRLKGDNEKHSRPNMGILGDLTLDLKDQRPDSHGLSPDSSENF	2100

Db 2041 ECISAMPKKPSRLKGDNEKHSFRNMGILGEDLTLDKDIQRPDSEHGLSPDSENF 2100  
Qy 2101 WFAIOEGANSIVSSIHOBAAACLRQASSDSILSKSGISLSPHLPDQBEKPT 2160  
Db 2101 WFAIOEGANSIVSSIHOBAAACLRQASSDSILSKSGISLSPHLPDQBEKPT 2160  
Qy 2161 SNKGRILKPGKSTLETKKIBESKGIKGGKVKYSLITGKVRNSISGQMQLQAN 2220  
Db 2161 SNKGRILKPGKSTLETKKIBESKGIKGGKVKYSLITGKVRNSISGQMQLQAN 2220  
Qy 2221 MFSISGRITMIHIGVRNSSSSTSVSKKGPPLKTPASKSPSEGTATTSRGAKPVS 2280  
Db 2221 MFSISGRITMIHIGVRNSSSSTSVSKKGPPLKTPASKSPSEGTATTSRGAKPVS 2280  
Qy 2281 ELSPVARTQISGSSKAPSRSGSDSTPSRPAQPLSRPIQSPGRNISPGRNISP 2340  
Db 2281 ELSPVARTQISGSSKAPSRSGSDSTPSRPAQPLSRPIQSPGRNISPGRNISP 2340  
Qy 2341 KLSQLPRTSSPTASTKSGSGKMYTSPGRMSQONLTQTGLSKNASSIPRESAS 2400  
Db 2341 KLSQLPRTSSPTASTKSGSGKMYTSPGRMSQONLTQTGLSKNASSIPRESAS 2400  
Qy 2401 LNMONGANGKVELSRMSTKSSGSDRSPVLRQSTFIKEAPSLRKLLESA 2460  
Db 2401 LNMONGANGKVELSRMSTKSSGSDRSPVLRQSTFIKEAPSLRKLLESA 2460  
Qy 2461 SPESLSPSRPASPTRSQAQTPVLSPLDMSLSTHSSVQAGMKLPPLNLSPTIEYND 2520  
Db 2461 SPESLSPSRPASPTRSQAQTPVLSPLDMSLSTHSSVQAGMKLPPLNLSPTIEYND 2520  
Qy 2521 RPAKHDIARSHSPSRPLPNRSCTWKREHSHSSLPVSTWRTGSSSILSAS 2580  
Db 2521 RPAKHDIARSHSPSRPLPNRSCTWKREHSHSSLPVSTWRTGSSSILSAS 2580  
Qy 2581 SEKASEDEKHVNSISGKQSKENOVSAKGTWRKIKENEFPSTNSTSTVSSGATNGA 2640  
Db 2581 SEKASEDEKHVNSISGKQSKENOVSAKGTWRKIKENEFPSTNSTSTVSSGATNGA 2640  
Qy 2641 KTLIIQMAPAVSKTEDVWRIEDCPINNPRSGRSTGNTPTVIDSVSEKANPNKSKDN 2700  
Db 2641 KTLIIQMAPAVSKTEDVWRIEDCPINNPRSGRSTGNTPTVIDSVSEKANPNKSKDN 2700  
Qy 2701 QAKQVNGVSPMRTVGLNRLNSFIQVADPKGTETKPGQNNPVVSETNESIVERT 2760  
Db 2701 QAKQVNGVSPMRTVGLNRLNSFIQVADPKGTETKPGQNNPVVSETNESIVERT 2760  
Qy 2761 PFSSSSSKHSPSGTVAARVTTPNPNPSRKSADSTARSQIPTPVNNTKRD 2820  
Db 2761 PFSSSSSKHSPSGTVAARVTTPNPNPSRKSADSTARSQIPTPVNNTKRD 2820  
Qy 2821 DSTESSGTQSPKRGSGVLTSTV 2843  
Db 2821 DSTESSGTQSPKRGSGVLTSTV 2843

RESULT 2

US-08-452-655B-7

Sequence 7, Application US/08452655B

Patent No. 5783666

GENERAL INFORMATION:

APPLICANT: ALBERTSEN, HANS  
APPLICANT: ANAND, RAKESH  
APPLICANT: CARLSON, MARY  
APPLICANT: GRODEN, JOANNA  
APPLICANT: HEDGE, PHILIP J.  
APPLICANT: JOSLYN, GEOFF  
APPLICANT: KINZLER, KENNETH  
APPLICANT: MARKHAM, ALEXANDER F.  
APPLICANT: NAKAMURA, YUSUKE  
APPLICANT: THLIVERIS, ANDREW  
TITLE OF INVENTION: INHERITED AND SOMATIC MUTATIONS OF APC  
TITLE OF INVENTION: GENE IN COLORECTAL CANCER IN HUMANS

NUMBER OF SEQUENCES: 102  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Banner & Witcoff, Ltd.  
STREET: 1001 G Street, NW  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20001-4598  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/452,655B  
FILING DATE: 25-MAY-1995  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/289,548  
FILING DATE: 12-AUG-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/741,940  
FILING DATE: 08-AUG-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Kagan, Sarah A.  
REGISTRATION NUMBER: 32,141  
REFERENCE/DOCKET NUMBER: 1107.49964  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-508-9100  
TELEFAX: 202-508-9299  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2843 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHEICAL: YES  
ANTI-SENSE: NO  
US-08-452-655B-7

Query Match 100.0%; Score 14566; DB 1; Length 2843;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 2843; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAAASYDQLLKQVEALKMENSILRQELDNHSLTKLETSANNKEVYLKQSGIDEAM 60  
Db 1 MAAASYDQLLKQVEALKMENSILRQELDNHSLTKLETSANNKEVYLKQSGIDEAM 60  
Qy 61 ASSQIDLLRLKELNLDSSNFPQVKLRKMSLRSGSREGSVSSRSGECSVPVMSFPR 120  
Db 61 ASSQIDLLRLKELNLDSSNFPQVKLRKMSLRSGSREGSVSSRSGECSVPVMSFPR 120  
Qy 121 RGFVNGSRESTGYLELEKERSILLADLDKEEKEKDWYQAQLQNLTKRIDSLPTENFSL 180  
Db 121 RGFVNGSRESTGYLELEKERSILLADLDKEEKEKDWYQAQLQNLTKRIDSLPTENFSL 180  
Qy 181 QDTMTREQLVEARQIRVAMEEQLGTCODMEKRAARRIARQIQIEKDLIRQLQSOAT 240  
Db 181 QDTMTREQLVEARQIRVAMEEQLGTCODMEKRAARRIARQIQIEKDLIRQLQSOAT 240  
Qy 241 EAERSQNKHETGSHDAERQNEQGVGEINMATSNGQGGSTTRMDHETASVLSSTHSA 300  
Db 241 EAERSQNKHETGSHDAERQNEQGVGEINMATSNGQGGSTTRMDHETASVLSSTHSA 300  
Qy 301 PRRLTSHLGTVMVYSLLSMLGTHDKDDMSRTLLAMSSQDSCISWRQSGCLPLLIQL 360  
Db 301 PRRLTSHLGTVMVYSLLSMLGTHDKDDMSRTLLAMSSQDSCISWRQSGCLPLLIQL 360  
Qy 361 HGNDKDSVLLGNRSGSKAARASAAHLNIIHQPDQKRRRIRVILHLEQIRAYCETC 420  
Db 361 HGNDKDSVLLGNRSGSKAARASAAHLNIIHQPDQKRRRIRVILHLEQIRAYCETC 420

QY 421 WEOEAHEPGVQDDKNPAPVHEHOICPAYCVLMKLSFOBEHRHAMNELGGLQAIABLLQ 480  
 Db 421 WEOEAHEPGVQDDKNPAPVHEHOICPAYCVLMKLSFDEHRHAMNELGGLQAIABLLQ 480  
 QY 481 VDEMYGLTNDHYSITLRRYAGMALTNLTFGDVANKATLCCKMGRALVAQLKSSSEDL 540  
 Db 481 VDEMYGLTNDHYSITLRRYAGMALTNLTFGDVANKATLCCKMGRALVAQLKSSSEDL 540  
 QY 541 QOVIASVLRNLSWRADVNSKKTIREVGSVKALMECALEVKKESTLKSVALNLSAHT 600  
 Db 541 QOVIASVLRNLSWRADVNSKKTIREVGSVKALMECALEVKKESTLKSVALNLSAHT 600  
 QY 601 ENKADICAVDGAFLAVGLTLYRSQNTNLAIIESGGILRNVSLLIATNEDHRQILRENN 660  
 Db 601 ENKADICAVDGAFLAVGLTLYRSQNTNLAIIESGGILRNVSLLIATNEDHRQILRENN 660  
 QY 661 CLQTLLOHLKSHSLTIVSNACGLTWNLSARNPKDOEALWDMGAVSMLKNLIHSHKHMIAM 720  
 Db 661 CLQTLLOHLKSHSLTIVSNACGLTWNLSARNPKDOEALWDMGAVSMLKNLIHSHKHMIAM 720  
 QY 721 GSAALRNLMANRPAPKVDANIMSPGSSLPFLVRKQKALEAELDAQHLSSETFDNIDNLS 780  
 Db 721 GSAALRNLMANRPAPKVDANIMSPGSSLPFLVRKQKALEAELDAQHLSSETFDNIDNLS 780  
 QY 781 PKASHRSKQKHQSLYGDYVFTNRHDDNRSDNFTGNMTVLPSPYLNNTVLPSSSSSRGS 840  
 Db 781 PKASHRSKQKHQSLYGDYVFTNRHDDNRSDNFTGNMTVLPSPYLNNTVLPSSSSSRGS 840  
 QY 841 LDSRSSEKDSLEBERGIGLGNVHPATENPGTSKGLQISTTAAOIAKVMEEVSAIHTS 900  
 Db 841 LDSRSSEKDSLEBERGIGLGNVHPATENPGTSKGLQISTTAAOIAKVMEEVSAIHTS 900  
 QY 901 QEDRSSGSTITLHCVTDERNALRRSSAAHSTNFTKSENRTCSMPYAKLEYKRSS 960  
 Db 901 QEDRSSGSTITLHCVTDERNALRRSSAAHSTNFTKSENRTCSMPYAKLEYKRSS 960  
 QY 961 NDSLNSVSSDGYCKRQOMKPSIESYSEDESFCYSQYPADLAHKIHSANHMDNDGE 1020  
 Db 961 NDSLNSVSSDGYCKRQOMKPSIESYSEDESFCYSQYPADLAHKIHSANHMDNDGE 1020  
 QY 1021 LDTPIYSLKYSDQLNSGRQSPONERWARPKHIIIEDEIKQSEQRNQSTTYPVYTE 1080  
 Db 1021 LDTPIYSLKYSDQLNSGRQSPONERWARPKHIIIEDEIKQSEQRNQSTTYPVYTE 1080  
 QY 1081 STDDXHLKFPQHFQOQECVSPYRSRGANGSETNRVGSNHGINQVQSOLCOEDDYEDDKP 1140  
 Db 1081 STDDXHLKFPQHFQOQECVSPYRSRGANGSETNRVGSNHGINQVQSOLCOEDDYEDDKP 1140  
 QY 1141 TNYSERVSEBEOHEBERPNYSIKYNEERKHVDOPIDYSLKYATDIPSSQKQSFPSKS 1200  
 Db 1141 TNYSERVSEBEOHEBERPNYSIKYNEERKHVDOPIDYSLKYATDIPSSQKQSFPSKS 1200  
 QY 1201 SSGQSSKTEHMSSESSTSTPSSNAKQNLHPSSAQSRSGQPOKAATCKVSSINQETIQ 1260  
 Db 1201 SSGQSSKTEHMSSESSTSTPSSNAKQNLHPSSAQSRSGQPOKAATCKVSSINQETIQ 1260  
 QY 1261 TYCVEDTPICFSRCSLSLSAABDEIGCNOQTCEADSNANTLOIAETKEKIGTRSAEDPV 1320  
 Db 1261 TYCVEDTPICFSRCSLSLSAABDEIGCNOQTCEADSNANTLOIAETKEKIGTRSAEDPV 1320  
 QY 1321 SEVPAVQHPRTKSRRLQSSLSSESARHKAFFSSGAKSPSKGAQTPKSPPEHYVQET 1380  
 Db 1321 SEVPAVQHPRTKSRRLQSSLSSESARHKAFFSSGAKSPSKGAQTPKSPPEHYVQET 1380  
 QY 1381 PLMFSRCTSVSSLDSPFSRGIASVQSPCSGVSGIISPSDLDPGQTMPPSRSKTTP 1440  
 Db 1381 PLMFSRCTSVSSLDSPFSRGIASVQSPCSGVSGIISPSDLDPGQTMPPSRSKTTP 1440  
 QY 1441 PPTQATQKEVPKNKAPTAKEKESGPKQAANVAQVQVLPDADTLHLHFAFESTPDGF 1500  
 Db 1441 PPTQATQKEVPKNKAPTAKEKESGPKQAANVAQVQVLPDADTLHLHFAFESTPDGF 1500  
 QY 1501 SCSSLSALSIDEPFIQKDVLELRIMPPVQENDNGNETESQPKESNENQSKAEKTIIDSE 1560

Db 1501 SCSSLSALSIDEPFIQKDVLELRIMPPVQENDNGNETESQPKESNENQSKAEKTIIDSE 1560  
 QY 1561 KOLLDDSDDDDIIELEBECIISAMPTKSRKAKKPAQTASKLPPPVARKPQOLPVYKLLPS 1620  
 Db 1561 KOLLDDSDDDDIIELEBECIISAMPTKSRKAKKPAQTASKLPPPVARKPQOLPVYKLLPS 1620  
 QY 1621 QNRLOPQKHVSFTPGDDMPRVYCEGTPINFSTATSLDITIESPPNLAAGEVGRGAQ 1680  
 Db 1621 QNRLOPQKHVSFTPGDDMPRVYCEGTPINFSTATSLDITIESPPNLAAGEVGRGAQ 1680  
 QY 1681 SGEFEKRTIIPTEGRSTDEAOGKTSVVTPELDDNKAEBGDIILAEICINSAMPKGSHPK 1740  
 Db 1681 SGEFEKRTIIPTEGRSTDEAOGKTSVVTPELDDNKAEBGDIILAEICINSAMPKGSHPK 1740  
 QY 1741 FRVKKIMDQVOQASASSAPNKNQLDGKKKPTSPVKPIPONTYRTRVRKNADSKNNLN 1800  
 Db 1741 FRVKKIMDQVOQASASSAPNKNQLDGKKKPTSPVKPIPONTYRTRVRKNADSKNNLN 1800  
 QY 1801 AERVFSDNKOSKQNLKXNSKDFNDKLPNNEDRVGSAFDSPHHYTPIEGTPYCFGRND 1860  
 Db 1801 AERVFSDNKOSKQNLKXNSKDFNDKLPNNEDRVGSAFDSPHHYTPIEGTPYCFGRND 1860  
 QY 1861 SLSLDFDDDDVDLSREKAELEKAKENKESAKVTSHTELTSNOOSANKTOAIKQPINR 1920  
 Db 1861 SLSLDFDDDDVDLSREKAELEKAKENKESAKVTSHTELTSNOOSANKTOAIKQPINR 1920  
 QY 1921 GQPKPILQKQSTFPQSSKDIIPDRGAATDEKLQNFPAIENTPVCFSHNSLSLSIDQENN 1980  
 Db 1921 GQPKPILQKQSTFPQSSKDIIPDRGAATDEKLQNFPAIENTPVCFSHNSLSLSIDQENN 1980  
 QY 1981 NKENEPIKETPPPOSQGEPPKQASGYAPKSFHVEDTPVCFSRNSLSLSIDSDDLLQ 2040  
 Db 1981 NKENEPIKETPPPOSQGEPPKQASGYAPKSFHVEDTPVCFSRNSLSLSIDSDDLLQ 2040  
 QY 2041 ECISASMPKKKPKRLKGDNEKHSFRNMGGITLGBDLTLDLKDIORPDSEHGLSPDSNFD 2100  
 Db 2041 ECISASMPKKKPKRLKGDNEKHSFRNMGGITLGBDLTLDLKDIORPDSEHGLSPDSNFD 2100  
 QY 2101 WKAIQEGANSIVSSLHQAAAAAALSRQASDSDSLKSGISGLSGSPFLTTPDOEKPF 2160  
 Db 2101 WKAIQEGANSIVSSLHQAAAAAALSRQASDSDSLKSGISGLSGSPFLTTPDOEKPF 2160  
 QY 2161 SNKGPRILLKPEKSTLETTKIESKGIKGGKVKYKSLITGKVRNSSEISQMKQPIQAN 2220  
 Db 2161 SNKGPRILLKPEKSTLETTKIESKGIKGGKVKYKSLITGKVRNSSEISQMKQPIQAN 2220  
 QY 2221 MFSISRGRTMIHIFGVNRNSSSTSPVSKGPKLTPASKSPSECOATTPRGAAPSVKS 2280  
 Db 2221 MFSISRGRTMIHIFGVNRNSSSTSPVSKGPKLTPASKSPSECOATTPRGAAPSVKS 2280  
 QY 2281 ELSPVAROTSOIGSSKAPSRSGRSDTSPRPAOPLSRPIQSPGRNSISPRNGISPPN 2340  
 Db 2281 ELSPVAROTSOIGSSKAPSRSGRSDTSPRPAOPLSRPIQSPGRNSISPRNGISPPN 2340  
 QY 2341 KLSQLPRTSSPSTASTKSSGSKMSYTSRQMSQNLTKQTGLSKNASSIPRSESASKG 2400  
 Db 2341 KLSQLPRTSSPSTASTKSSGSKMSYTSRQMSQNLTKQTGLSKNASSIPRSESASKG 2400  
 QY 2401 LNMNNGANGANKVELSRMSSTKSSGESDRSEPRVLVROSTFIKEAPSPTLRKLBSA 2460  
 Db 2401 LNMNNGANGANKVELSRMSSTKSSGESDRSEPRVLVROSTFIKEAPSPTLRKLBSA 2460  
 QY 2461 SPESLSPSSRPASPTRSOAQTPVLSPLDMSLSTHSSVQAGWRKLPPLNLSPIEYNDG 2520  
 Db 2461 SPESLSPSSRPASPTRSOAQTPVLSPLDMSLSTHSSVQAGWRKLPPLNLSPIEYNDG 2520  
 QY 2521 RPAKHDHIAHSHSSEPSPLPINRSGTWKREHSHKSSSLPRVSTWRTGSSSSILASSES 2580  
 Db 2521 RPAKHDHIAHSHSSEPSPLPINRSGTWKREHSHKSSSLPRVSTWRTGSSSSILASSES 2580  
 QY 2581 SEKAKSEDEKHNSTSGTKQSKENQVSAKGTWRKIKENEFSTNSTGOTVSSSGATNGAES 2640



Db 2581 SEKASEDEKHVNSISGTHQSKENQVSAKGTWRKTKENEFSPTNSTQTSVSGATNGA8S 2640  
Qy 2641 KTLIYQMAPAVSKTBDVWVRIEDCPINNPRSGRSGPTGNTPPVIDSVSEKAMPNIKDSKDN 2700  
Db 2641 KTLIYQMAPAVSKTBDVWVRIEDCPINNPRSGRSGPTGNTPPVIDSVSEKAMPNIKDSKDN 2700  
Qy 2701 QAKQNVGSGVPMRTVGLNRLNSFIQVDAPQKTEIKPGQNNPVPVSETNESSIVERT 2760  
Db 2701 QAKQNVGSGVPMRTVGLNRLNSFIQVDAPQKTEIKPGQNNPVPVSETNESSIVERT 2760  
Qy 2761 PFSSSSSKHSSPSTGAARVTFPNYNSPRKSSADTSARPSQIPTPVNNTKKRDSKT 2820  
Db 2761 PFSSSSSKHSSPSTGAARVTFPNYNSPRKSSADTSARPSQIPTPVNNTKKRDSKT 2820  
Qy 2821 DSTSSSGTQSPXRHSGSYLVTSV 2843  
Db 2821 DSTSSSGTQSPXRHSGSYLVTSV 2843

RESULT 3  
US-08-450-582-2  
; Sequence 2, Application US/08450582  
; Patent No. 6114124  
; GENERAL INFORMATION:  
; APPLICANT: ALBERTSEN, HANS  
; APPLICANT: ANAND, RAKESH  
; APPLICANT: CARLSON, MARY  
; APPLICANT: GRODEN, JOANNA  
; APPLICANT: HEDGE, PHILIP J.  
; APPLICANT: JOSLYN, GEOFF  
; APPLICANT: KENZLER, KENNETH  
; APPLICANT: MARKHAM, ALEXANDER F.  
; APPLICANT: NAKAMURA, YUSUKE  
; APPLICANT: THLIVERIS, ANDREW  
; TITLE OF INVENTION: INHERITED AND SOMATIC MUTATIONS OF APC  
; TITLE OF INVENTION: GENE IN COLORECTAL CANCER IN HUMANS  
; NUMBER OF SEQUENCES: 102  
; CORRESPONDENCE ADDRESS:  
; ADDRESSER: Banner & Witcoff, Ltd.  
; STREET: 1001 G Street, NW  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20001-4598  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/450,582  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/452,655  
; FILING DATE: 25-MAY-1995  
; APPLICATION NUMBER: US 08/289,548  
; FILING DATE: 12-AUG-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/741,940  
; FILING DATE: 08-AUG-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Kagan, Sarah A.  
; REGISTRATION NUMBER: 32,141  
; REFERENCE/DOCKET NUMBER: 1107.49964  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-508-9100  
; TELEFAX: 202-508-9299  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2843 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear

; MOLECULE TYPE: protein  
US-08-450-582-2  
Query Match 100.0%; Score 14566; DB 3; Length 2843;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2843; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 MAASVYDQLLKQVEALQKENSINLRQELDESNHLTKLETSANMKEVLKQLOGSIEDEAM 60  
Db 1 MAASVYDQLLKQVEALQKENSINLRQELDESNHLTKLETSANMKEVLKQLOGSIEDEAM 60  
Qy 61 ASSQIDLLRLKELNLDSSNFFGVKLRKWSLRSGVSGREGSVSSRSGSGSPVPMGSFPR 120  
Db 61 ASSQIDLLRLKELNLDSSNFFGVKLRKWSLRSGVSGREGSVSSRSGSGSPVPMGSFPR 120  
Qy 121 RGFVNGSRESTGYLELEKERSLLADLDKEEKEKDWYAAQLQNLTKRIDSLPLTENFSL 180  
Db 121 RGFVNGSRESTGYLELEKERSLLADLDKEEKEKDWYAAQLQNLTKRIDSLPLTENFSL 180  
Qy 181 QDTMTREOLEYEAQIRVAVEEOLGTCDMEKQAORRIARIQOIEKDILRIQLLOSOAT 240  
Db 181 QDTMTREOLEYEAQIRVAVEEOLGTCDMEKQAORRIARIQOIEKDILRIQLLOSOAT 240  
Qy 241 EAERSSQNKHETGSHDAERQNEGQGVGEINMATSGNGQSGSTTRMDHETASVLSSSSTHSA 300  
Db 241 EAERSSQNKHETGSHDAERQNEGQGVGEINMATSGNGQSGSTTRMDHETASVLSSSSTHSA 300  
Qy 301 PRRLTSHLGTKEVMYISLLSMLGTHDKDDMSRILLAMSSQSDSCI SMRQSGCPLLIQLL 360  
Db 301 PRRLTSHLGTKEVMYISLLSMLGTHDKDDMSRILLAMSSQSDSCI SMRQSGCPLLIQLL 360  
Qy 361 HGNDKDSVLLGNRSGSKEARASAAALHNTIHSOPDDKRGRRIRVHLLEQIRAYCETC 420  
Db 361 HGNDKDSVLLGNRSGSKEARASAAALHNTIHSOPDDKRGRRIRVHLLEQIRAYCETC 420  
Qy 421 WEMQEAHEPGMDQKNPMPAPVEHQICPAVCVLMKLSFDEEHRHAMNELGGLQAIABLLQ 480  
Db 421 WEMQEAHEPGMDQKNPMPAPVEHQICPAVCVLMKLSFDEEHRHAMNELGGLQAIABLLQ 480  
Qy 481 VDCMYGLTNDHYSITLRRYAGWALTNLTFDGVANKATLCMSKGMCPALVAQLKSEEDL 540  
Db 481 VDCMYGLTNDHYSITLRRYAGWALTNLTFDGVANKATLCMSKGMCPALVAQLKSEEDL 540  
Qy 541 QQVIASVLRNLSWRADVNSKTLREVGSKVAMECALVKKESTLKSLSVLSALWNLSAHCT 600  
Db 541 QQVIASVLRNLSWRADVNSKTLREVGSKVAMECALVKKESTLKSLSVLSALWNLSAHCT 600  
Qy 601 ENKADICAVDGAFLVGTLYRSQNTNLAIIESGGIIRNVSSLIATNEDHRIQLRENN 660  
Db 601 ENKADICAVDGAFLVGTLYRSQNTNLAIIESGGIIRNVSSLIATNEDHRIQLRENN 660  
Qy 661 CLQTLLOHLKSHSLTI VSNACGTLNLSARNPKQDEALWDMGAVSMLKNLIHSGKHMIAM 720  
Db 661 CLQTLLOHLKSHSLTI VSNACGTLNLSARNPKQDEALWDMGAVSMLKNLIHSGKHMIAM 720  
Qy 721 GSAALNLMANRPAPKYKDNIMSPGSSLSLHVHRKOKALEABLDQHLSETFDNIDNLS 780  
Db 721 GSAALNLMANRPAPKYKDNIMSPGSSLSLHVHRKOKALEABLDQHLSETFDNIDNLS 780  
Qy 781 PKASHRSKORHKOSLYGDYVFDNRHDDNDSNFNTGNMTVLSPYNTTTLVLPSSSSSRGS 840  
Db 781 PKASHRSKORHKOSLYGDYVFDNRHDDNDSNFNTGNMTVLSPYNTTTLVLPSSSSSRGS 840  
Qy 841 LDSSRSKDRSLRERGIGLGNYPATENPGTSSKRGLOISTTAAQIAKVMEEVSAIHTS 900  
Db 841 LDSSRSKDRSLRERGIGLGNYPATENPGTSSKRGLOISTTAAQIAKVMEEVSAIHTS 900  
Qy 901 QEDRSSGTTTELHCVTDERNALRRSSAAHSTHNTYFTKSENGNRCTCSMPYAKLEYKRSS 960  
Db 901 QEDRSSGTTTELHCVTDERNALRRSSAAHSTHNTYFTKSENGNRCTCSMPYAKLEYKRSS 960  
Qy 961 NDSLNSVSSSDGYGKRGOMKPSIESYEDDESXFCSYGOYPADLAHKIHSANHMDNDGE 1020

Db 961 NDSLNSVSSDGYCKRQMKPSTESYEDDESKECSYGOYPADLAHKIHSANHWDDNDGE 1020  
Qy 1021 LDTPINTSLKYSDQLNSGRQSPQNERWARPKHIIIEDEIKQSEQRNQSTYPVYTE 1080  
Db 1021 LDTPINTSLKYSDQLNSGRQSPQNERWARPKHIIIEDEIKQSEQRNQSTYPVYTE 1080  
Qy 1081 STDDKHLKFOHFGQOCVSPYSRGANGSETNRVGSNHGINOVQSOLCOEDDYEDDKP 1140  
Db 1081 STDDKHLKFOHFGQOCVSPYSRGANGSETNRVGSNHGINOVQSOLCOEDDYEDDKP 1140  
Qy 1141 TNSERYSEOEHEBERPTNYSIKYNEEKHRVDQPIDYSLKYATDIPSSQKQSFSSKS 1200  
Db 1141 TNSERYSEOEHEBERPTNYSIKYNEEKHRVDQPIDYSLKYATDIPSSQKQSFSSKS 1200  
Qy 1201 SSGOSSKTEHMSSESSENTSPSSNAKRONQHLHSSAQSRSQOPKAATCKVSSINGETIQ 1260  
Db 1201 SSGOSSKTEHMSSESSENTSPSSNAKRONQHLHSSAQSRSQOPKAATCKVSSINGETIQ 1260  
Qy 1261 TYCVEDTPICFSRCSLSLSAEDDEIGCNOTTQEADSAANTLOIAEIKKIGTRSAEDPV 1320  
Db 1261 TYCVEDTPICFSRCSLSLSAEDDEIGCNOTTQEADSAANTLOIAEIKKIGTRSAEDPV 1320  
Qy 1321 SEVPAVQCHPRTKSRLQGSLSSESARHKAFFSSGAKSPSKGAQTPKSPPEHYVQET 1380  
Db 1321 SEVPAVQCHPRTKSRLQGSLSSESARHKAFFSSGAKSPSKGAQTPKSPPEHYVQET 1380  
Qy 1381 PLMFSRCTSVSSLDSPESRSITASSVQPCSGVSGIISPSDLPDPGOTMPPSRKTPP 1440  
Db 1381 PLMFSRCTSVSSLDSPESRSITASSVQPCSGVSGIISPSDLPDPGOTMPPSRKTPP 1440  
Qy 1441 PPPQTAQTKREVPKNKAAPTAKRESGPKQAANVAORVQVLPDADTLHFAFESTPDGF 1500  
Db 1441 PPPQTAQTKREVPKNKAAPTAKRESGPKQAANVAORVQVLPDADTLHFAFESTPDGF 1500  
Qy 1501 SCSSLSALSLDEPFIQKVELRIMPVQENDNGNETESQPKESNENKEAEKTIIDSE 1560  
Db 1501 SCSSLSALSLDEPFIQKVELRIMPVQENDNGNETESQPKESNENKEAEKTIIDSE 1560  
Qy 1561 KDLLDDDDDDDEILEECIIISAMPTKSSRKAKKPAQTAASKLPPPVARKPSQLPVYKLLPS 1620  
Db 1561 KDLLDDDDDDDEILEECIIISAMPTKSSRKAKKPAQTAASKLPPPVARKPSQLPVYKLLPS 1620  
Qy 1621 QNRLOPKHVSFTPGDDMPRVYCEVETPIINFSTATSLSDLTIESPNELAAAGVGGQAQ 1680  
Db 1621 QNRLOPKHVSFTPGDDMPRVYCEVETPIINFSTATSLSDLTIESPNELAAAGVGGQAQ 1680  
Qy 1681 SGEFEKDTTPTGERSTDEAQQGKTSSVTIPELDDNKAEBGDIILAEICINSAMPKGSHKP 1740  
Db 1681 SGEFEKDTTPTGERSTDEAQQGKTSSVTIPELDDNKAEBGDIILAEICINSAMPKGSHKP 1740  
Qy 1741 FRVKIMQVQOQASASSAPNKNQLDGKKKKTPSPVKPIPQNTYEYTRVRKNADSKNNLN 1800  
Db 1741 FRVKIMQVQOQASASSAPNKNQLDGKKKKTPSPVKPIPQNTYEYTRVRKNADSKNNLN 1800  
Qy 1801 AERFSDNKSCKONLKNNSKDFNDKLPNNEDVRGSAFDSPHHYTPIEGTFCFSRND 1860  
Db 1801 AERFSDNKSCKONLKNNSKDFNDKLPNNEDVRGSAFDSPHHYTPIEGTFCFSRND 1860  
Qy 1861 SLSSLDLDDDDVDLSREKAEELKAKENKESAEKVTSHTELTSNQOQANKTQAIAKQPINR 1920  
Db 1861 SLSSLDLDDDDVDLSREKAEELKAKENKESAEKVTSHTELTSNQOQANKTQAIAKQPINR 1920  
Qy 1921 GQPKPILOKOSTPQSSKQIPDRGAANDKLONFALIENTPVCFSHNSLSLSLDIDQENN 1980  
Db 1921 GQPKPILOKOSTPQSSKQIPDRGAANDKLONFALIENTPVCFSHNSLSLSLDIDQENN 1980  
Qy 1981 NKNEPIKETEPDPSQGEPSKQASGYAPKS FVVEDTFCVCFSRNSSLISLSDSEDDLLQ 2040  
Db 1981 NKNEPIKETEPDPSQGEPSKQASGYAPKS FVVEDTFCVCFSRNSSLISLSDSEDDLLQ 2040  
Qy 2041 ECISAMPKXKPSRLKGDNEKSPRNMGILGEDTLIDIKTORPDSHGLSPDSENFD 2100  
Db 2041 ECISAMPKXKPSRLKGDNEKSPRNMGILGEDTLIDIKTORPDSHGLSPDSENFD 2100

Qy 2101 WKATOEGANSIVSSILHQAACILSRQASSDSILSLKSGISLGSPHLTPDOEEKPFT 2160  
Db 2101 WKATOEGANSIVSSILHQAACILSRQASSDSILSLKSGISLGSPHLTPDOEEKPFT 2160  
Qy 2161 SNKGPRILLKPEKSTLETKKIETESSKGIKGGKVVYKSLITGKVSNSSEISQOMKQPLQAN 2220  
Db 2161 SNKGPRILLKPEKSTLETKKIETESSKGIKGGKVVYKSLITGKVSNSSEISQOMKQPLQAN 2220  
Qy 2221 MPSTISRGTMHIHFGVRNNSSTSPVKKGPPLKTPASKSPSECOQTATTSPRGAKPSVKS 2280  
Db 2221 MPSTISRGTMHIHFGVRNNSSTSPVKKGPPLKTPASKSPSECOQTATTSPRGAKPSVKS 2280  
Qy 2281 ELSVPVARTSOIGSSSKAPSRSGSRDSTPSPAPQPLSRPIQSPGRNSISPGRNGISPPN 2340  
Db 2281 ELSVPVARTSOIGSSSKAPSRSGSRDSTPSPAPQPLSRPIQSPGRNSISPGRNGISPPN 2340  
Qy 2341 KLSQLPRTSSPSTASTKSSGSKVSYTSPGRMQSQNLTKOTGLSKNASSIPRSESASKG 2400  
Db 2341 KLSQLPRTSSPSTASTKSSGSKVSYTSPGRMQSQNLTKOTGLSKNASSIPRSESASKG 2400  
Qy 2401 LQNMNNGANGANKVELSRMSSTKSSGSESDSERPVLVROSTFIKEAPSPTLRKLRESA 2460  
Db 2401 LQNMNNGANGANKVELSRMSSTKSSGSESDSERPVLVROSTFIKEAPSPTLRKLRESA 2460  
Qy 2461 SPESLSPSSRRPASPTRSQOATPVLSPSLPDMSLSTHSSVQAGGWKLPNLSPTIEYNDG 2520  
Db 2461 SPESLSPSSRRPASPTRSQOATPVLSPSLPDMSLSTHSSVQAGGWKLPNLSPTIEYNDG 2520  
Qy 2521 RPAKHDTARSHSESPRLPINRSGTWKREHSHKSSSLPRVSTWRTGSSSSILSASSES 2580  
Db 2521 RPAKHDTARSHSESPRLPINRSGTWKREHSHKSSSLPRVSTWRTGSSSSILSASSES 2580  
Qy 2581 SEKAKSEDEKHVNSISGTKQSKENQVSAKGTWRIKENERFSTNSTQTVSSGATNAGAES 2640  
Db 2581 SEKAKSEDEKHVNSISGTKQSKENQVSAKGTWRIKENERFSTNSTQTVSSGATNAGAES 2640  
Qy 2641 KTLIYQMAPAVSKTEDVWVRIEDCPINNPRSGRPTGNTPPVIDSVSEKANPNIKDSKDN 2700  
Db 2641 KTLIYQMAPAVSKTEDVWVRIEDCPINNPRSGRPTGNTPPVIDSVSEKANPNIKDSKDN 2700  
Qy 2701 QAKQVNGVSGVPMRTVGLNRLNSFIQVADAPDQKGTIKNQONNPVVPVSETNESSIVERT 2760  
Db 2701 QAKQVNGVSGVPMRTVGLNRLNSFIQVADAPDQKGTIKNQONNPVVPVSETNESSIVERT 2760  
Qy 2761 PFSSSSSKHSSPGTVAARVTPFNYPNPSPKSSADSTSARPSQIPTPVNNNTKKRDSKT 2820  
Db 2761 PFSSSSSKHSSPGTVAARVTPFNYPNPSPKSSADSTSARPSQIPTPVNNNTKKRDSKT 2820  
Qy 2821 DSTESSGTSQSPKRGSGVYLTSV 2843  
Db 2821 DSTESSGTSQSPKRGSGVYLTSV 2843

## RESULT 4

US-08-450-582-7

; Sequence 7, Application US/08450582

; Patent No. 6114124

## GENERAL INFORMATION:

; APPLICANT: ALBERTSEN, HANS

; APPLICANT: ANAND, RAKESH

; APPLICANT: CARLSON, MARY

; APPLICANT: GRODEN, JOANNA

; APPLICANT: HEDGE, PHILIP J.

; APPLICANT: JOSLYN, GEOFF

; APPLICANT: KINZLER, KENNETH

; APPLICANT: MARKHAM, ALEXANDER F.

; APPLICANT: NAKAMURA, YUSUKE

; APPLICANT: THLIVERIS, ANDREW

; TITLE OF INVENTION: INHERITED AND SOMATIC MUTATIONS OF APC

; NUMBER OF SEQUENCES: 102

; CORRESPONDENCE ADDRESS:

ADDRESSEE: Banner & Witcoff, Ltd.  
STREET: 1001 G Street, NW  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20001-4598  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/450,582  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/452,655  
FILING DATE: 25-MAY-1995  
APPLICATION NUMBER: US 08/289,548  
FILING DATE: 12-AUG-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/741,940  
FILING DATE: 08-AUG-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Kagan, Sarah A.  
REGISTRATION NUMBER: 32,141  
REFERENCE/DOCKET NUMBER: 1107.49964  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-508-9100  
TELEFAX: 202-508-9299  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2843 amino acids  
TYPE: amino acid  
STRADEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: YES  
ANTI-SENSE: NO  
US-08-450-582-7

Query Match 100.0%; Score 14566; DB 3; Length 2843;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 2843; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MAAAYDQLLKQVEALKMNSLRQELNDSNHLTKLEASNMKEVLKQLOGSIEDAM	60
DB	1	MAAAYDQLLKQVEALKMNSLRQELNDSNHLTKLEASNMKEVLKQLOGSIEDAM	60
QY	61	ASSGQIDLLERKELNLDSSNPPGVKLRSMKLSRSGREGSVSRSGECSPVPMGSPFR	120
DB	61	ASSGQIDLLERKELNLDSSNPPGVKLRSMKLSRSGREGSVSRSGECSPVPMGSPFR	120
QY	121	RGFVNGSRESTGYLEBEKERSILLADLDKEKEKDWYVAQLNLTKRIDSILPTENFSL	180
DB	121	RGFVNGSRESTGYLEBEKERSILLADLDKEKEKDWYVAQLNLTKRIDSILPTENFSL	180
QY	181	QDTMTROLEYEARQIRVAMEEQLCTCDMEKRAORRIARICQIEKDILIRQLQSOAT	240
DB	181	QDTMTROLEYEARQIRVAMEEQLCTCDMEKRAORRIARICQIEKDILIRQLQSOAT	240
QY	241	EAEERSQKHETGSHDAERQNGQGVGEINMATSGNGQGSTTRMDHETASVLSSSSTHSA	300
DB	241	EAEERSQKHETGSHDAERQNGQGVGEINMATSGNGQGSTTRMDHETASVLSSSSTHSA	300
QY	301	PRRLTSHLGTQVEMVYSLISMLGTHDKDMSRTLLAMSSSQSDSCISMROSGCLPILIQLL	360
DB	301	PRRLTSHLGTQVEMVYSLISMLGTHDKDMSRTLLAMSSSQSDSCISMROSGCLPILIQLL	360
QY	361	HGNDKDSVLLGNRSGKEARASAAALHNIHSPQDDKGRREIRVHLLEQIRAYCETC	420
DB	361	HGNDKDSVLLGNRSGKEARASAAALHNIHSPQDDKGRREIRVHLLEQIRAYCETC	420

QY	421	WEWQEAHEPGMDQDNKMPAPVEHQICPAVCVLMKLSFDEHRHAMNELGGQIAIELIQ	480
DB	421	WEWQEAHEPGMDQDNKMPAPVEHQICPAVCVLMKLSFDEHRHAMNELGGQIAIELIQ	480
QY	481	VDCEMYGLTNDHYSITLRRYAGMALTNLTFGDVANKATILCSMKGCRALVAOLKSSIDL	540
DB	481	VDCEMYGLTNDHYSITLRRYAGMALTNLTFGDVANKATILCSMKGCRALVAOLKSSIDL	540
QY	541	QOVIASVLRNLNLSWRADVNSKTLREVGSVKALMECALEVKEKSTLSKLSALWLSAHT	600
DB	541	QOVIASVLRNLNLSWRADVNSKTLREVGSVKALMECALEVKEKSTLSKLSALWLSAHT	600
QY	601	ENKADICAVDGAFLVGLTYRSQNTLAIIBSGGILRNVSLLIATNEDHQIILRENN	660
DB	601	ENKADICAVDGAFLVGLTYRSQNTLAIIBSGGILRNVSLLIATNEDHQIILRENN	660
QY	661	CLQTLQHLKSHSLTIVSNACGLTNLSARNPKQDALMDGAVSMLKNLHSHKHMIAM	720
DB	661	CLQTLQHLKSHSLTIVSNACGLTNLSARNPKQDALMDGAVSMLKNLHSHKHMIAM	720
QY	721	GSAALRNLMANRPAPKYKDIANIMSPGSSLPSSHVRKQKALEAELDAQHLSSETFDNLS	780
DB	721	GSAALRNLMANRPAPKYKDIANIMSPGSSLPSSHVRKQKALEAELDAQHLSSETFDNLS	780
QY	781	PKASHRSKQHKOSLYGDYVDFDNRHDDNRSDNFNTGNMTVLSPLYNTTTLVLPSSSSSRGS	840
DB	781	PKASHRSKQHKOSLYGDYVDFDNRHDDNRSDNFNTGNMTVLSPLYNTTTLVLPSSSSSRGS	840
QY	841	LDSSRGKDRSLERERIGLGNYPATENPGTSSKRGLOISTTAAQIAKWEVSVAIHTS	900
DB	841	LDSSRGKDRSLERERIGLGNYPATENPGTSSKRGLOISTTAAQIAKWEVSVAIHTS	900
QY	901	QEDRSSGTTTLCVTDERNALRRSAAHSTNYFTKSENRRCTCSMPYAKLEYKRSS	960
DB	901	QEDRSSGTTTLCVTDERNALRRSAAHSTNYFTKSENRRCTCSMPYAKLEYKRSS	960
QY	961	NDLSNVSSSDGYGKRGQMKPSIESYSEDDSEKFCYGYQYPADLAHKKIHSANEMDNDGE	1020
DB	961	NDLSNVSSSDGYGKRGQMKPSIESYSEDDSEKFCYGYQYPADLAHKKIHSANEMDNDGE	1020
QY	1021	LDTPINYSKYSDQNLNSGRQSPQNERWAPKHIIIEDEIKQSEQRQSRNQSTTPYVTE	1080
DB	1021	LDTPINYSKYSDQNLNSGRQSPQNERWAPKHIIIEDEIKQSEQRQSRNQSTTPYVTE	1080
QY	1081	STDDXHLKFPQHFQGCVCVPSYRGANGSETNRVGSNHCINQNVQSCLCQEDDYEDDKP	1140
DB	1081	STDDXHLKFPQHFQGCVCVPSYRGANGSETNRVGSNHCINQNVQSCLCQEDDYEDDKP	1140
QY	1141	TNYSERYSEEEQHEEERPTNYSIKYNEEKRVHDQPIDYSLKYATDIPSSQKQSFSPKS	1200
DB	1141	TNYSERYSEEEQHEEERPTNYSIKYNEEKRVHDQPIDYSLKYATDIPSSQKQSFSPKS	1200
QY	1201	SSGSSKTEHMSSESSENTSTPSSNAKRONLHPSSAQSRSGOPKAACTCKVSSINQETIQ	1260
DB	1201	SSGSSKTEHMSSESSENTSTPSSNAKRONLHPSSAQSRSGOPKAACTCKVSSINQETIQ	1260
QY	1261	TYCVEDTPICFSRCSLSLSAEDIEGCTQTTQEADSAANTLQIAEIKKIGTRSAEDPV	1320
DB	1261	TYCVEDTPICFSRCSLSLSAEDIEGCTQTTQEADSAANTLQIAEIKKIGTRSAEDPV	1320
QY	1321	SEVPASVQHPRTKSSRLQSSLSSESARHKAVFFSSGAKSPSKSGAQTPKSPPEHYVQET	1380
DB	1321	SEVPASVQHPRTKSSRLQSSLSSESARHKAVFFSSGAKSPSKSGAQTPKSPPEHYVQET	1380
QY	1381	PLMFSTRCTSVSSLDSPESRSIASSVQSEPCSGMVSGIISPSDLPSDPSQWMPSPSRKTPP	1440
DB	1381	PLMFSTRCTSVSSLDSPESRSIASSVQSEPCSGMVSGIISPSDLPSDPSQWMPSPSRKTPP	1440
QY	1441	PPPTAQTKREVPKNKAPTAKESGPKQAANAAVORVQVLPDADTLLHFATESTPDGF	1500
DB	1441	PPPTAQTKREVPKNKAPTAKESGPKQAANAAVORVQVLPDADTLLHFATESTPDGF	1500
QY	1501	SCSSLSALSILDEFFIKQDVELRIMPVQENDNGNETESQPKESNENQKEAEKTIIDSE	1560

Db 1501 SCSSLSALSILDEPFIQKDELIMPPVQENDNGNETESQPKESNEQKEAEKIDSE 1560  
 QY 1561 KOLLDDDDDDIEILECIIISAMPTKSRRAKPAQTASKLPPVAKPSQOLPVYKLLPS 1620  
 Db 1561 KOLLDDDDDDIEILECIIISAMPTKSRRAKPAQTASKLPPVAKPSQOLPVYKLLPS 1620  
 QY 1621 QNRLOPKHVSFTPGDDMPVYCEVETPINFSTATSLDITIESPPNELAAGGVGGGAQ 1680  
 Db 1621 QNRLOPKHVSFTPGDDMPVYCEVETPINFSTATSLDITIESPPNELAAGGVGGGAQ 1680  
 QY 1681 SGEPEKDDTIPTTEGRSTDEAAGGKTSSVTIPELDDNKAEGDILAEICINSAMPKGSHPK 1740  
 Db 1681 SGEPEKDDTIPTTEGRSTDEAAGGKTSSVTIPELDDNKAEGDILAEICINSAMPKGSHPK 1740  
 QY 1741 FRVKKIMDOVQOASASSAFNKNQDGGKKKPTSPVKPIPONTETRYTRVRKNADSKNNLN 1800  
 Db 1741 FRVKKIMDOVQOASASSAFNKNQDGGKKKPTSPVKPIPONTETRYTRVRKNADSKNNLN 1800  
 QY 1801 AERFSDNKKONLXNSKDFNDKLPNNEDVRGSAFDPSPHHYPTIEGTPYCFSRND 1860  
 Db 1801 AERFSDNKKONLXNSKDFNDKLPNNEDVRGSAFDPSPHHYPTIEGTPYCFSRND 1860  
 QY 1861 SLSSLDFFDDVDLSREKAEIRAKENKESAEKVTSHTELTSSQOSANKTOATAKOPINR 1920  
 Db 1861 SLSSLDFFDDVDLSREKAEIRAKENKESAEKVTSHTELTSSQOSANKTOATAKOPINR 1920  
 QY 1921 GQPKPILQKSTFPQSSKDIPTDGAATDEKLFQFAIENTPVCFSHNSLSLSDIDQENN 1980  
 Db 1921 GQPKPILQKSTFPQSSKDIPTDGAATDEKLFQFAIENTPVCFSHNSLSLSDIDQENN 1980  
 QY 1981 NKNEPTIKETEPDQSGEPKQASGAPKSFHVEDTPVCFSRNSLSLSDIDSEDDLIQ 2040  
 Db 1981 NKNEPTIKETEPDQSGEPKQASGAPKSFHVEDTPVCFSRNSLSLSDIDSEDDLIQ 2040  
 QY 2041 ECISAMPKPKKSRSLKDNKESPRNMGGILGEDITLDLKDQRPDSEHGLSPDSENF 2100  
 Db 2041 ECISAMPKPKKSRSLKDNKESPRNMGGILGEDITLDLKDQRPDSEHGLSPDSENF 2100  
 QY 2101 WKAIQEGANSIVSLHQAACLSROASDSISLSKSGISLGSFPFHLTPDQEKPT 2160  
 Db 2101 WKAIQEGANSIVSLHQAACLSROASDSISLSKSGISLGSFPFHLTPDQEKPT 2160  
 QY 2161 SNKGPRILKPEKSTLTETKIESEKIGKGVYKSLITGKVRNSSEISGQWKQPLQAN 2220  
 Db 2161 SNKGPRILKPEKSTLTETKIESEKIGKGVYKSLITGKVRNSSEISGQWKQPLQAN 2220  
 QY 2221 MPSISRGRTHIIPGVNRSSSTSPVSKGPPKTPASKSPSGQTATTSRGAKEPSVK 2280  
 Db 2221 MPSISRGRTHIIPGVNRSSSTSPVSKGPPKTPASKSPSGQTATTSRGAKEPSVK 2280  
 QY 2281 ELSPVARTQSIGSSKAPSRGSRDSTPRPAQPLSRPIQSPGNSISPGENGISPN 2340  
 Db 2281 ELSPVARTQSIGSSKAPSRGSRDSTPRPAQPLSRPIQSPGNSISPGENGISPN 2340  
 QY 2341 KLSQLPRTSPSTASTKSSSGRMSYTSPOQMSQONLTKQGLSKNASSIPRSESASG 2400  
 Db 2341 KLSQLPRTSPSTASTKSSSGRMSYTSPOQMSQONLTKQGLSKNASSIPRSESASG 2400  
 QY 2401 LQWNTNGANKKVELSRMSTKSSGSDRSPRPVLRQSTFIKAPSPFTLRKLEESA 2460  
 Db 2401 LQWNTNGANKKVELSRMSTKSSGSDRSPRPVLRQSTFIKAPSPFTLRKLEESA 2460  
 QY 2461 SFESLSPSSRPASPTASQATPVLSPLPMSLTHSSVQAGWRKLPNLSPTIYNDG 2520  
 Db 2461 SFESLSPSSRPASPTASQATPVLSPLPMSLTHSSVQAGWRKLPNLSPTIYNDG 2520  
 QY 2521 RPAKRHDIAHSGESPSRLPINRSGTWKREHSHSSSLPRVSTWRTGSSSILSASS 2580  
 Db 2521 RPAKRHDIAHSGESPSRLPINRSGTWKREHSHSSSLPRVSTWRTGSSSILSASS 2580  
 QY 2581 SEKAKSEDEKHVNSIGTKQSKENQVSAKGTWRKIKENEFSPNTNSTQTVSSGATNGAES 2640

RESULT 5  
 US-08-821-355A-7  
 ; Sequence 7, Application US/08821355A  
 ; Patent No. 5851775  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Barker, Nick  
 ; APPLICANT: Clevers, Hans  
 ; APPLICANT: Korinek, Vladimir  
 ; APPLICANT: Morin, Patrice  
 ; APPLICANT: Kinzler, Kenneth  
 ; APPLICANT: Vogelstein, Bert  
 ; APPLICANT: Sparks, Andrew  
 ; TITLE OF INVENTION: Beta Catenin, TCF-4, and APC  
 ; TITLE OF INVENTION: Interact to Prevent Cancer  
 ; NUMBER OF SEQUENCES: 11  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Banner & Witcoff, Ltd.  
 ; STREET: 1001 G Street, N.W.  
 ; CITY: Washington  
 ; STATE: DC  
 ; COUNTRY: USA  
 ; ZIP: 20001  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Diskette  
 ; COMPUTER: IBM Compatible  
 ; OPERATING SYSTEM: DOS  
 ; SOFTWARE: FASECQ for Windows Version 2.0  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/821,355A  
 ; FILING DATE: 20-MAR-1997  
 ; CLASSIFICATION: 514  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER:  
 ; FILING DATE:  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Kagan, Sarah A  
 ; REGISTRATION NUMBER: 32,145  
 ; REFERENCE/DOCKET NUMBER: 1107.05064  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 202-508-9100  
 ; TELEFAX: 202-508-9299  
 ; TELEX: 97430 BMB UT  
 ; INFORMATION FOR SEQ ID NO: 7:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 2973 amino acids  
 ; TYPE: amino acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: No. 5851775e  
 ; US-08-821-355A-7

Query Match 100.0%; Score 14566; DB 2; Length 2973;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 2843; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAASDYDOLLKQVLEALKNWENLNROLEDNEDNHLTKLETEASNKEVLKOLQGSIEDEAM 60  
 Db 1 MAASDYDOLLKQVLEALKNWENLNROLEDNEDNHLTKLETEASNKEVLKOLQGSIEDEAM 60  
 QY 61 ASSGOIDILLERLKEINLDSNPPGVKLRKMSLSYSGREGSVSRSGECSPVPMGSPFR 120  
 Db 61 ASSGOIDILLERLKEINLDSNPPGVKLRKMSLSYSGREGSVSRSGECSPVPMGSPFR 120  
 QY 121 RGFVNGSRESTGYLELEKERSLILADLDKEKEKOWYIACQNLTKRIDSLPLTENFSL 180  
 Db 121 RGFVNGSRESTGYLELEKERSLILADLDKEKEKOWYIACQNLTKRIDSLPLTENFSL 180  
 QY 181 QDTRRQLEYEARQIRVAMEEQLGTCODMEKRAQRRIARIQOIEKOLIRIQLQSOAT 240  
 Db 181 QDTRRQLEYEARQIRVAMEEQLGTCODMEKRAQRRIARIQOIEKOLIRIQLQSOAT 240  
 QY 241 EARSSONKHETGSHDAERONEGOVGEINNATSGNGOGSTTRMDHETASVLSSSSTHSA 300  
 Db 241 EARSSONKHETGSHDAERONEGOVGEINNATSGNGOGSTTRMDHETASVLSSSSTHSA 300  
 QY 301 PRRLTSHLGTKEVYMSLLSMLGTHDKDDMSRTLLAMSSQSDSCTSMFQSCCLPLLIQLL 360  
 Db 301 PRRLTSHLGTKEVYMSLLSMLGTHDKDDMSRTLLAMSSQSDSCTSMFQSCCLPLLIQLL 360  
 QY 361 HGNDKDSVLLNGSRGSKARAPASAAHLNIIHSQPDYGRREIRVLHLLQIRAYCETC 420  
 Db 361 HGNDKDSVLLNGSRGSKARAPASAAHLNIIHSQPDYGRREIRVLHLLQIRAYCETC 420  
 QY 421 WEQBAHBPBGMDQKQMPAPVEHOICPAVCVLMKLSFDEEHRHAMNELGLQIAIELLO 480  
 Db 421 WEQBAHBPBGMDQKQMPAPVEHOICPAVCVLMKLSFDEEHRHAMNELGLQIAIELLO 480  
 QY 481 VDCEMYGLTNDHYSITLRYAGMALTNLTFGDVANKATLCNKGCMRALVAQLKSESEDL 540  
 Db 481 VDCEMYGLTNDHYSITLRYAGMALTNLTFGDVANKATLCNKGCMRALVAQLKSESEDL 540  
 QY 541 QOVIASVLNLSWRADVNSKTLREVGSKALMECALEVKESTLKSVLKALWNLASHCT 600  
 Db 541 QOVIASVLNLSWRADVNSKTLREVGSKALMECALEVKESTLKSVLKALWNLASHCT 600  
 QY 601 ENKADICAVDAGALAFVGLTLYRSQNTLAIIESGGGLRNVSSLIATNEDHROILRENN 660  
 Db 601 ENKADICAVDAGALAFVGLTLYRSQNTLAIIESGGGLRNVSSLIATNEDHROILRENN 660  
 QY 661 CLQTLLOHLKSHSLTI VSNACGTLNLSARNPKDQKALWDMGAVSMLKNLTHSKHMTAM 720  
 Db 661 CLQTLLOHLKSHSLTI VSNACGTLNLSARNPKDQKALWDMGAVSMLKNLTHSKHMTAM 720  
 QY 721 GSAALRLNLMANRPKYKDANTMSPGSSLP SLHVRKQKALAEALDAQHLSETFDNIDLS 780  
 Db 721 GSAALRLNLMANRPKYKDANTMSPGSSLP SLHVRKQKALAEALDAQHLSETFDNIDLS 780  
 QY 781 PXASHRSKORHKOSLYGDYVFTNRHDDNRSNFTGNMTVLSPYLNTTVLPSSSSSRGS 840  
 Db 781 PXASHRSKORHKOSLYGDYVFTNRHDDNRSNFTGNMTVLSPYLNTTVLPSSSSSRGS 840  
 QY 841 LDSSRSKDRSLEREGTGLGNYPHATENPGTSSKRGLOISTTAQIAKWEEVSAIHTS 900  
 Db 841 LDSSRSKDRSLEREGTGLGNYPHATENPGTSSKRGLOISTTAQIAKWEEVSAIHTS 900  
 QY 901 QDRSSSGSTTELCVTDERNALRRSSAAHTNFTYNTFKSENSNRCTSMYPKALEYKRS 960  
 Db 901 QDRSSSGSTTELCVTDERNALRRSSAAHTNFTYNTFKSENSNRCTSMYPKALEYKRS 960  
 QY 961 NDSLNSVSSSDGYKRGQKQKPSIESYSEDDSKFCSYGYPADLAHKHHSANHMDNDGE 1020  
 Db 961 NDSLNSVSSSDGYKRGQKQKPSIESYSEDDSKFCSYGYPADLAHKHHSANHMDNDGE 1020  
 QY 1021 LUTPYNLSKYDQELNSGRQSPQNERWARKPHIIEDEIKQSEORQORQSNQSTTYPVYTE 1080  
 Db 1021 LUTPYNLSKYDQELNSGRQSPQNERWARKPHIIEDEIKQSEORQORQSNQSTTYPVYTE 1080

QY 1081 STDDKHLKFPQHFQOQECVSPYRSRGANGSETNVRVGSNHEGINQNVQSLSLQBEDDYEDDKP 1140  
 Db 1081 STDDKHLKFPQHFQOQECVSPYRSRGANGSETNVRVGSNHEGINQNVQSLSLQBEDDYEDDKP 1140  
 QY 1141 TNYSEYRSEEEQHEBERPTNYSIKYNEEKRVHDQIDIDYSLKYATDIPSSOKOSFSPKS 1200  
 Db 1141 TNYSEYRSEEEQHEBERPTNYSIKYNEEKRVHDQIDIDYSLKYATDIPSSOKOSFSPKS 1200  
 QY 1201 SSCQSKTEHMSSSSNTSTPSSNAKRNQQLHPSSAQSRSGQPQKAATCKVSSINQETIQ 1260  
 Db 1201 SSCQSKTEHMSSSSNTSTPSSNAKRNQQLHPSSAQSRSGQPQKAATCKVSSINQETIQ 1260  
 QY 1261 TYCVETDTPCFRCSLSSLSAEDBIGCQNTQOEADSNANTLOIAIEKIKIGTRSAEDPV 1320  
 Db 1261 TYCVETDTPCFRCSLSSLSAEDBIGCQNTQOEADSNANTLOIAIEKIKIGTRSAEDPV 1320  
 QY 1321 SEVPANVSQHPRTKSSRLQGSLSSESARHKAVEFFSSCAKSPSKSGAQTPKSPPEHYVQET 1380  
 Db 1321 SEVPANVSQHPRTKSSRLQGSLSSESARHKAVEFFSSCAKSPSKSGAQTPKSPPEHYVQET 1380  
 QY 1381 PLMFSRCTSVSSLDSTFESRSIASSVQSEPCSGMVSGIISPSDLDPSPGQTMPPSRKSTPP 1440  
 Db 1381 PLMFSRCTSVSSLDSTFESRSIASSVQSEPCSGMVSGIISPSDLDPSPGQTMPPSRKSTPP 1440  
 QY 1441 PPQTAQTKREVPKNKAAPTAEKEESGPKOAAVNAAYQVRQVLPDADTLHLHFAESTPDGF 1500  
 Db 1441 PPQTAQTKREVPKNKAAPTAEKEESGPKOAAVNAAYQVRQVLPDADTLHLHFAESTPDGF 1500  
 QY 1501 SCSSLSALSALDPEFFIQKOVLEIRIMPVQENDNGNETESEQPKESNENQKEAEKIDSE 1560  
 Db 1501 SCSSLSALSALDPEFFIQKOVLEIRIMPVQENDNGNETESEQPKESNENQKEAEKIDSE 1560  
 QY 1561 KDLDDSDDDDIIEBECIISAMPTKSSKAKPAOTASKLPPPVARKSOLPVYKLLPS 1620  
 Db 1561 KDLDDSDDDDIIEBECIISAMPTKSSKAKPAOTASKLPPPVARKSOLPVYKLLPS 1620  
 QY 1621 QNRLOPKHVSFTPGDDMPRVYCVETGTFNFTSTATSLDITIESPPNELAAGSGVRGGAQ 1680  
 Db 1621 QNRLOPKHVSFTPGDDMPRVYCVETGTFNFTSTATSLDITIESPPNELAAGSGVRGGAQ 1680  
 QY 1681 SGFEKEDTTPTEGRSTDEAQQGKTSSVTPELDDNKAEBGDILACINSAMPKGSHP 1740  
 Db 1681 SGFEKEDTTPTEGRSTDEAQQGKTSSVTPELDDNKAEBGDILACINSAMPKGSHP 1740  
 QY 1741 FRVKKIMQVQOASASSAPKNQKGGKKTSPVKPIPONTEYTRVRKNADSKNNLN 1800  
 Db 1741 FRVKKIMQVQOASASSAPKNQKGGKKTSPVKPIPONTEYTRVRKNADSKNNLN 1800  
 QY 1801 AERFSDNKDCKKQNLKNNSKDFNDKLPNNEDVRGSGFAPDSPHHYTPIEGTYPYCFSRND 1860  
 Db 1801 AERFSDNKDCKKQNLKNNSKDFNDKLPNNEDVRGSGFAPDSPHHYTPIEGTYPYCFSRND 1860  
 QY 1861 SLSSLDDEDDDDVLSREKAEELRKAENKSEAKVTSHTELTSNQOANKTOATAKOPINR 1920  
 Db 1861 SLSSLDDEDDDDVLSREKAEELRKAENKSEAKVTSHTELTSNQOANKTOATAKOPINR 1920  
 QY 1921 GQPKPILQKQSTFPQSSKQIPDRGAATDEKLQNFATIENTPVCFSHNSLSLSDIDQENN 1980  
 Db 1921 GQPKPILQKQSTFPQSSKQIPDRGAATDEKLQNFATIENTPVCFSHNSLSLSDIDQENN 1980  
 QY 1981 NKNEPTEKTEPPDSQGEPSKPOAGYAPKSFVEDTTPVCFSNSLSLSIDSEDDLLQ 2040  
 Db 1981 NKNEPTEKTEPPDSQGEPSKPOAGYAPKSFVEDTTPVCFSNSLSLSIDSEDDLLQ 2040  
 QY 2041 ECTSSAMPKPKKPSRLKGDNEKHSPRNMGGILGEDTLTLDKIQRPDSEHGLSPDSENF 2100  
 Db 2041 ECTSSAMPKPKKPSRLKGDNEKHSPRNMGGILGEDTLTLDKIQRPDSEHGLSPDSENF 2100  
 QY 2101 WKATQEGANSIVSLHAAAAACLSRQASSDSLSLSKSGISLGSPFHLTPQOEKPF 2160  
 Db 2101 WKATQEGANSIVSLHAAAAACLSRQASSDSLSLSKSGISLGSPFHLTPQOEKPF 2160  
 QY 2161 SNKGPRILKQEKSTLTETKLIBSESKGKGGKVKYKSLITGKVRNSNIEISGQMKQKQAN 2220

2161	Db	SNKGRPLRKPEKXSTLTETKKBESKGIKGGKVYKSLITGKVRNSEISGQWKQPLQAN	2220
2221	Qy	MPISIRGRRTMIHIPGVNNSSTSPVSKGPPPLKTPASKSPSEGQTATTPRGAFTSVKS	2280
2221	Db	MPISIRGRRTMIHIPGVNNSSTSPVSKGPPPLKTPASKSPSEGQTATTPRGAFTSVKS	2280
2281	Qy	ELSPVARQTSQICGSSKAPSRSGRSDTPRRPAQOPLSRPIQSPGRNSISPGRNGISPPN	2340
2281	Db	ELSPVARQTSQICGSSKAPSRSGRSDTPRRPAQOPLSRPIQSPGRNSISPGRNGISPPN	2340
2341	Qy	KLSQLPRTSPSTASTKSSGSKMSYTPSPGRQMSQQNLTKQTGLSKNASSIPRSESASKG	2400
2341	Db	KLSQLPRTSPSTASTKSSGSKMSYTPSPGRQMSQQNLTKQTGLSKNASSIPRSESASKG	2400
2401	Qy	LNOMNNGNGANKKVEUSRMSTKSSGSESPSRERPVLVROSTFIKEAPSPTLRKLKLEESA	2460
2401	Db	LNOMNNGNGANKKVEUSRMSTKSSGSESPSRERPVLVROSTFIKEAPSPTLRKLKLEESA	2460
2461	Qy	SPFESLSPSSRPASPTRSOAQTPVLSPELSPDWLSLTHSSVQAGGWRKLPPNLSPTIIBYNDG	2520
2461	Db	SPFESLSPSSRPASPTRSOAQTPVLSPELSPDWLSLTHSSVQAGGWRKLPPNLSPTIIBYNDG	2520
2521	Qy	RPAKRHDIKAESHESPERLPINBSGWTWRKRBHSHKSSSLPRVSTWRKTGSSSTLSASSES	2580
2521	Db	RPAKRHDIKAESHESPERLPINBSGWTWRKRBHSHKSSSLPRVSTWRKTGSSSTLSASSES	2580
2581	Qy	SEKAKSEDEKXVANSISGTQKSKENQVSAKGTWKIKENEPSPTNSTSQTVSSGATNGAES	2640
2581	Db	SEKAKSEDEKXVANSISGTQKSKENQVSAKGTWKIKENEPSPTNSTSQTVSSGATNGAES	2640
2641	Qy	KTLIYQMAPAVSKTDDVWVIEOCPINNPBSGRSPTCGNTPPVLDSDVSEKANPINIKDOKN	2700
2641	Db	KTLIYQMAPAVSKTDDVWVIEOCPINNPBSGRSPTCGNTPPVLDSDVSEKANPINIKDOKN	2700
2701	Qy	QAKQNVNGSGVPMRTVGLENLRNLSFIQVADPADQKGTIEIKPQNNPVPVSETNESSIVERT	2760
2701	Db	QAKQNVNGSGVPMRTVGLENLRNLSFIQVADPADQKGTIEIKPQNNPVPVSETNESSIVERT	2760
2761	Qy	PFSSSSSKHSSPSGTVAAEVTPFNPNPSPRKSSADSTSARPSQITPVPVNNNTKKRDSKT	2820
2761	Db	PFSSSSSKHSSPSGTVAAEVTPFNPNPSPRKSSADSTSARPSQITPVPVNNNTKKRDSKT	2820
2821	Qy	DSTESSGTSQPKRHSGSYLTVSV	2843
2821	Db	DSTESSGTSQPKRHSGSYLTVSV	2843

```

; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ For Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/003,687A
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/821,355
; FILING DATE: 20-MAR-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Kagan, Sarah A
; REGISTRATION NUMBER: 32,145
; REFERENCE/DOCKET NUMBER: 1107.05064
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-508-9100
; TELEFAX: 202-508-9299
; TELEX: 97430 BMB UT
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2973 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: No. 5998600e
US-09-003-687A-7

Query Match 100.0%; Score 14566; DB 2; Length 2973;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2843; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAAASYDQLLKQVEALKMENSNLRLQLEDNSNHLTKLETEASNWKVLLKQLGSIIDEAM 60
Db 1 MAAASYDQLLKQVEALKMENSNLRLQLEDNSNHLTKLETEASNWKVLLKQLGSIIDEAM 60

Qy 61 ASSGQIDLRLKELNLDSSNFFGVKLRKWSLRSYGSREGSVSSRSGCSPVPMGSPFR 120
Db 61 ASSGQIDLRLKELNLDSSNFFGVKLRKWSLRSYGSREGSVSSRSGCSPVPMGSPFR 120

Qy 121 RGFVNSRGSTGYLELEKERSLLADLKEEKKDWYQAQNLTKRIDSPLTENFSL 180
Db 121 RGFVNSRGSTGYLELEKERSLLADLKEEKKDWYQAQNLTKRIDSPLTENFSL 180

Qy 181 QDTMTRRQLEYEARQIRVAMEEQLGTCDQWEKRAQRRIARIQQIEKDILIRLQLQSQT 240
Db 181 QDTMTRRQLEYEARQIRVAMEEQLGTCDQWEKRAQRRIARIQQIEKDILIRLQLQSQT 240

Qy 241 EARRSQNHETGSHDAERQNEQGVGEINMATSGNGCGSTTMDHETASVLSSSSTHSA 300
Db 241 EARRSQNHETGSHDAERQNEQGVGEINMATSGNGCGSTTMDHETASVLSSSSTHSA 300

Qy 301 PRRLTSHLGTKEVMYVLSLLSMGLTDPDKDMSRTLLAMSSSQSDCISMRQSGCIPLLIQLL 360
Db 301 PRRLTSHLGTKEVMYVLSLLSMGLTDPDKDMSRTLLAMSSSQSDCISMRQSGCIPLLIQLL 360

Qy 361 HGNDKDSVLGNRSGKKEARASAAALHNIHSQPDQKGRREIRVHLLEQIRAYCETC 420
Db 361 HGNDKDSVLGNRSGKKEARASAAALHNIHSQPDQKGRREIRVHLLEQIRAYCETC 420

Qy 421 WEQEAHEPGMDQKNPMPAPVEHQICPAVCVLMKLSFDEEHRHAMNELGGHQAIAELLQ 480
Db 421 WEQEAHEPGMDQKNPMPAPVEHQICPAVCVLMKLSFDEEHRHAMNELGGHQAIAELLQ 480

Qy 481 VDCEMYGLTNDHYSITLRRYAGMALTNLTFGDVANKATLCSMKGCWRAIVAQLKSESDL 540
Db 481 VDCEMYGLTNDHYSITLRRYAGMALTNLTFGDVANKATLCSMKGCWRAIVAQLKSESDL 540

Qy 541 QQVIASVLRNLSWRADVNSKKTILREYGSVKALMECALEVKEKSTLKSVLASLWNLSAHCT 600
Db 541 QQVIASVLRNLSWRADVNSKKTILREYGSVKALMECALEVKEKSTLKSVLASLWNLSAHCT 600

Qy 601 ENKADICAVDGAFLVGLTLYRSQNTLTAIESGGGILRNYSSTIATNEDHRQILRENN 660
Db 601 ENKADICAVDGAFLVGLTLYRSQNTLTAIESGGGILRNYSSTIATNEDHRQILRENN 660

```

QY 661 CLQTLLOHLKSHSLTIVSNACGTLWNLGARNPKDOEALWDMGAVSMLKNLILHSHKMIAM 720  
DB 661 CLQTLLOHLKSHSLTIVSNACGTLWNLGARNPKDOEALWDMGAVSMLKNLILHSHKMIAM 720  
QY 721 GSAALRNLMANRPARYKXANIMSPGSSLPSSLHVRKQXALAEALDAQHLSETFFONIDNLS 780  
DB 721 GSAALRNLMANRPARYKXANIMSPGSSLPSSLHVRKQXALAEALDAQHLSETFFONIDNLS 780  
QY 781 PKASHRSKORHQSLYGDYVFTDNRHDDNRSNENTGNMTVLSPYLNTTTLVPSSSSRGS 840  
DB 781 PKASHRSKORHQSLYGDYVFTDNRHDDNRSNENTGNMTVLSPYLNTTTLVPSSSSRGS 840  
QY 841 LPSRSEKDRSLERERIGLGNYPATENPGTSSKRGLOISTTAAQIAKWEEVSIAHTS 900  
DB 841 LPSRSEKDRSLERERIGLGNYPATENPGTSSKRGLOISTTAAQIAKWEEVSIAHTS 900  
QY 901 QEDRSGSTTELHCVTDERNALRRSAATHNTNFTKSENSNETCSMPYAKLEYKRSS 960  
DB 901 QEDRSGSTTELHCVTDERNALRRSAATHNTNFTKSENSNETCSMPYAKLEYKRSS 960  
QY 961 NDSLNSVSSDGYGKRGQMKPSIESYSEDDSKFCYGOYPADLAHKIHSANHMDDNDGE 1020  
DB 961 NDSLNSVSSDGYGKRGQMKPSIESYSEDDSKFCYGOYPADLAHKIHSANHMDDNDGE 1020  
QY 1021 LPTPINYSLKYSDEQLNSGRQSPQNERWARPKHIIEDBIKQSEORQSRNOSTTYPVYTE 1080  
DB 1021 LPTPINYSLKYSDEQLNSGRQSPQNERWARPKHIIEDBIKQSEORQSRNOSTTYPVYTE 1080  
QY 1081 STDDKHLKFPQHFQOQECVSPYRSGANGSENRVGNHGINQVNSQSLCOEDDYEDDKP 1140  
DB 1081 STDDKHLKFPQHFQOQECVSPYRSGANGSENRVGNHGINQVNSQSLCOEDDYEDDKP 1140  
QY 1141 TNYSERYSEEBEHEERPTNYSIIKYNEBKRVDPIDYSLKYATDIPSSQKQSFSEK 1200  
DB 1141 TNYSERYSEEBEHEERPTNYSIIKYNEBKRVDPIDYSLKYATDIPSSQKQSFSEK 1200  
QY 1201 SSGQSKTEHMSSENTSTPSSNAGRNQOLHPSSAQSRSGOPQKAAATCKVSSINQETIQ 1260  
DB 1201 SSGQSKTEHMSSENTSTPSSNAGRNQOLHPSSAQSRSGOPQKAAATCKVSSINQETIQ 1260  
QY 1261 TYCVSDTPTCFGRCSLSLSAEDEIGCNQTOQADSAANTQIAEIKEIKIGTRSAEDPV 1320  
DB 1261 TYCVSDTPTCFGRCSLSLSAEDEIGCNQTOQADSAANTQIAEIKEIKIGTRSAEDPV 1320  
QY 1321 SEVPASVQHPRTKSRLOQSSLSSESARHKAVEFSSGAKSPSKSGAOTPKPPEHYQET 1380  
DB 1321 SEVPASVQHPRTKSRLOQSSLSSESARHKAVEFSSGAKSPSKSGAOTPKPPEHYQET 1380  
QY 1381 PLMFGRCTSVSLDGFESRSIASSVQSEPCQWVGIIISPSDLPSQGTMPPSRSKTPP 1440  
DB 1381 PLMFGRCTSVSLDGFESRSIASSVQSEPCQWVGIIISPSDLPSQGTMPPSRSKTPP 1440  
QY 1441 PPPQTAQTKREVPKNAKPTAEKRESGPKQAAVNAAVQRVQVLPDADTLHLFATESTPDGF 1500  
DB 1441 PPPQTAQTKREVPKNAKPTAEKRESGPKQAAVNAAVQRVQVLPDADTLHLFATESTPDGF 1500  
QY 1501 SCSSLSALSLEPFIQKDVLRIMPPVQENDNGNETSEQPKESNENQKEAEXTIDSE 1560  
DB 1501 SCSSLSALSLEPFIQKDVLRIMPPVQENDNGNETSEQPKESNENQKEAEXTIDSE 1560  
QY 1561 KOLLDDSDDDDIIEILEECIIISAMPTKSRKAKKPAQTASKLPPPVARPKSQLPVYKLLPS 1620  
DB 1561 KOLLDDSDDDDIIEILEECIIISAMPTKSRKAKKPAQTASKLPPPVARPKSQLPVYKLLPS 1620  
QY 1621 QNRLOPKHVSFTPGDDMPRVYCVETGPIINFSTATSLDLTIESPNELAAEGVRRGAQ 1680  
DB 1621 QNRLOPKHVSFTPGDDMPRVYCVETGPIINFSTATSLDLTIESPNELAAEGVRRGAQ 1680  
QY 1681 SGEFEKRDITPTEGRSTDEAQGGKTSVTPPELDDNKAEEGDIILAEICINSAMPKGSHPK 1740  
DB 1681 SGEFEKRDITPTEGRSTDEAQGGKTSVTPPELDDNKAEEGDIILAEICINSAMPKGSHPK 1740

QY 1741 FRVKKIMDOVQOASASSAPNKNQLDGKKKPTSPVKPIPQNTFYRTRVKNADSKNNLN 1800  
DB 1741 FRVKKIMDOVQOASASSAPNKNQLDGKKKPTSPVKPIPQNTFYRTRVKNADSKNNLN 1800  
QY 1801 AERVFSNKNDSKQNLKNNSKDFNDKLPNNEDVRGSPAFDSPHHYTPIEGTYCFCSRND 1860  
DB 1801 AERVFSNKNDSKQNLKNNSKDFNDKLPNNEDVRGSPAFDSPHHYTPIEGTYCFCSRND 1860  
QY 1861 SLGSLDEDDDDVLSREKAEALRKAENKESAEKVTSHTLTSTNOQANKTQALAKQPINR 1920  
DB 1861 SLGSLDEDDDDVLSREKAEALRKAENKESAEKVTSHTLTSTNOQANKTQALAKQPINR 1920  
QY 1921 GQPKPILQKQSTFPQSSKDIIPDRGAATDEKLQNAFAIENTPVCFSHNSLSLSLSDIDQENN 1980  
DB 1921 GQPKPILQKQSTFPQSSKDIIPDRGAATDEKLQNAFAIENTPVCFSHNSLSLSLSDIDQENN 1980  
QY 1981 NKNEPIKETEPDPDSQGEPSKPOASGVAPKSFHVEDTTPVCFSRNSLSLSLSIUSEDLLQ 2040  
DB 1981 NKNEPIKETEPDPDSQGEPSKPOASGVAPKSFHVEDTTPVCFSRNSLSLSLSIUSEDLLQ 2040  
QY 2041 ECISAMPKPKKPSRLKGDNEKESPRNMGGILGEDTLTLDKIQRPDSEHGLSPDSENF 2100  
DB 2041 ECISAMPKPKKPSRLKGDNEKESPRNMGGILGEDTLTLDKIQRPDSEHGLSPDSENF 2100  
QY 2101 WKAIQEGANSIVSSLHQAASAAACLSRQASDSDSIILSKSGISLGSPPHLTPDOEEKPFT 2160  
DB 2101 WKAIQEGANSIVSSLHQAASAAACLSRQASDSDSIILSKSGISLGSPPHLTPDOEEKPFT 2160  
QY 2161 SNKGPRILKFGKSTLTETKKIESESGIKGKKVYKSLITGKVRNSSEISGQMKQPLQAN 2220  
DB 2161 SNKGPRILKFGKSTLTETKKIESESGIKGKKVYKSLITGKVRNSSEISGQMKQPLQAN 2220  
QY 2221 MPSISRORTMIHIPGVNNSSTSPVSKGPPKPTASKPSQEGOTATTSPRAKPSVKS 2280  
DB 2221 MPSISRORTMIHIPGVNNSSTSPVSKGPPKPTASKPSQEGOTATTSPRAKPSVKS 2280  
QY 2281 ELSVARSQTSQIGSSKAPSRSGSRDTPSRPAQOPLSRPIQSPGRNSISFGRANGISPPN 2340  
DB 2281 ELSVARSQTSQIGSSKAPSRSGSRDTPSRPAQOPLSRPIQSPGRNSISFGRANGISPPN 2340  
QY 2341 KLSQLPRTSPSTASTKSSGSGKMYTSPGRQMSQQNLTKQTLGSKNASSIPRSESASK 2400  
DB 2341 KLSQLPRTSPSTASTKSSGSGKMYTSPGRQMSQQNLTKQTLGSKNASSIPRSESASK 2400  
QY 2401 LNQMNNGANKKVELSEMSSTKSSGSESDSRSPRVLRQSTTIKEAPSPTLARKLEESA 2460  
DB 2401 LNQMNNGANKKVELSEMSSTKSSGSESDSRSPRVLRQSTTIKEAPSPTLARKLEESA 2460  
QY 2461 SFESLSPSRPASPTRSQAOTPVLSPLPDMVSLSTHSSVQAGWRKLPPLNLSPTIEYNDG 2520  
DB 2461 SFESLSPSRPASPTRSQAOTPVLSPLPDMVSLSTHSSVQAGWRKLPPLNLSPTIEYNDG 2520  
QY 2521 RPAKRHDIARSHSESPRLPINRSGTWKREHSHKSSLPVSTVWRTTGGSSSILSASSES 2580  
DB 2521 RPAKRHDIARSHSESPRLPINRSGTWKREHSHKSSLPVSTVWRTTGGSSSILSASSES 2580  
QY 2581 SEKAKSEDEKHVNSISGTKOSKENQVSAKTWIKIENEFSPNTSQTSSGATNGAES 2640  
DB 2581 SEKAKSEDEKHVNSISGTKOSKENQVSAKTWIKIENEFSPNTSQTSSGATNGAES 2640  
QY 2641 KTLIYQMAPAVSKTDEVVVRIEDCPIINNPRSGSPGTGNTPPVIDSVSEKANPNIKDSKON 2700  
DB 2641 KTLIYQMAPAVSKTDEVVVRIEDCPIINNPRSGSPGTGNTPPVIDSVSEKANPNIKDSKON 2700  
QY 2701 QAKONVNGSVPMRTVGLNRLNSFTQVADPQDQTEIKPQONNPVPSSTNESSIVERT 2760  
DB 2701 QAKONVNGSVPMRTVGLNRLNSFTQVADPQDQTEIKPQONNPVPSSTNESSIVERT 2760  
QY 2761 PFSSSSSSSKHSPSGTVAARVTFNPNPSPRKSSADSTSARPSQIPTFVNNTTKKDSKT 2820  
DB 2761 PFSSSSSSSKHSPSGTVAARVTFNPNPSPRKSSADSTSARPSQIPTFVNNTTKKDSKT 2820  
QY 2821 DSTESSGTQSPKRHSGSYLITSV 2843





Db 2761 PFSSSSSKHSPSGTVAARVTPFNPNPGRKSSADSTARPSPQIPTPVNNNTKRDST 2820

Qy 2821 DSTESSGTQSPKRHSGLYLTSV 2843

Db 2821 DSTESSGTQSPKRHSGLYLTSV 2843

RESULT 8

US-07-741-940-7

; Sequence 7, Application US/07741940

; Patent No. 5352775

; GENERAL INFORMATION:

; APPLICANT: ALBERTSEN, HANS

; APPLICANT: ANAND, RAKESH

; APPLICANT: CARLSON, MARY

; APPLICANT: GRODEN, JOANNA

; APPLICANT: HEDGE, PHILIP J.

; APPLICANT: JOSLYN, GEOFF

; APPLICANT: KINZLER, KENNETH

; APPLICANT: MARKHAM, ALEXANDER F.

; APPLICANT: NAKAMURA, YUSUKE

; APPLICANT: THIVIERIS, ANDREW

; TITLE OF INVENTION: INHERITED AND SOMATIC MUTATIONS OF APC

; TITLE OF INVENTION: GENE IN COLORECTAL CANCER IN HUMANS

; NUMBER OF SEQUENCES: 94

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Banner, Birch, McKie & Beckett

; STREET: 1001 G Street, NW

; CITY: Washington

; STATE: D.C.

; COUNTRY: USA

; ZIP: 20001-4598

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent In Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/07/741,940

; FILING DATE: 19920109

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: Kagan, Sarah A.

; REGISTRATION NUMBER: 32,141

; REFERENCE/DOCKET NUMBER: 1107.035574

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 202-508-9100

; TELEFAX: 202-508-9299

; INFORMATION FOR SEQ ID NO: 7:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 2842 amino acids

; TYPE: AMINO ACID

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; ORIGINAL SOURCE:

; ORGANISM: Homo sapiens

; IMMEDIATE SOURCE:

; CLONE: APC

US-07-741-940-7

Query Match 99.9%; Score 14548.5; DB 1; Length 2842;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 2842; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

Qy 1 MAASVDLLKQVEALKMENSRLRQLEDNSNHLTKLETSANMKVYLKQSGSIDAM 60

Db 1 MAASVDLLKQVEALKMENSRLRQLEDNSNHLTKLETSANMKVYLKQSGSIDAM 60

Qy 61 ASSGQIDLLRLKELNLDSSNFFGVKLRKQWLSRYSRSGSVSSRSGECSVPVPMGSFPR 120

Db 61 ASSGQIDLLRLKELNLDSSNFFGVKLRKQWLSRYSRSGSVSSRSGECSVPVPMGSFPR 120

Db 1681 SGEFEKRDITPTEGRSTDEAQGKTSVVTIPELDDNKAEEDGILAEICINSAMPKSHKP 1740

Qy 1741 FRVKIMQVQOQASASSAPNKNQDGGKKKTPSVKPIPONTEYRTRVRKNADSKNLN 1800

Db 1741 FRVKIMQVQOQASASSAPNKNQDGGKKKTPSVKPIPONTEYRTRVRKNADSKNLN 1800

Qy 1801 AERFSDNKKSKQNLKNNKDFNDKLPNNEDRVGSAFDPSPHHYPIEGTPYCFSEND 1860

Db 1801 AERFSDNKKSKQNLKNNKDFNDKLPNNEDRVGSAFDPSPHHYPIEGTPYCFSEND 1860

Qy 1861 SLSSLDPDODDVLDSREKAEKAKENKSEAKYTSHTELTSNQQSANKTQAIKQPINR 1920

Db 1861 SLSSLDPDODDVLDSREKAEKAKENKSEAKYTSHTELTSNQQSANKTQAIKQPINR 1920

Qy 1921 GQPKILQKSTFFPOSSKDIIDRGAAATDEKLNFAIENTPVCFSHNSLSLSIDIDQNN 1980

Db 1921 GQPKILQKSTFFPOSSKDIIDRGAAATDEKLNFAIENTPVCFSHNSLSLSIDIDQNN 1980

Qy 1981 NKENPIKETEPDPSQGPSPQAGYAPKSPFHVEDTTPVCFSRNSSLSSIDSEDDLLQ 2040

Db 1981 NKENPIKETEPDPSQGPSPQAGYAPKSPFHVEDTTPVCFSRNSSLSSIDSEDDLLQ 2040

Qy 2041 ECISAMPKKKPSRLKGDNEKHSRNMGGILGEDLTLDLKDIOQPDSEHGLSPDSENF 2100

Db 2041 ECISAMPKKKPSRLKGDNEKHSRNMGGILGEDLTLDLKDIOQPDSEHGLSPDSENF 2100

Qy 2101 WKAIQEGANSIVSSLHQAAAACLSRQASPSDSILSLKSGISLSPHLPDQEKPT 2160

Db 2101 WKAIQEGANSIVSSLHQAAAACLSRQASPSDSILSLKSGISLSPHLPDQEKPT 2160

Qy 2161 SNKGRILKPKBKSTLETKLISESGIKGKKVYKSLITCKVRNSLSISGMQKPLQAN 2220

Db 2161 SNKGRILKPKBKSTLETKLISESGIKGKKVYKSLITCKVRNSLSISGMQKPLQAN 2220

Qy 2221 MFSISGRRTMIHPGVNRSSSTSPVSKGKPLKTPASKSPSEGTATTSRPGAKPSVKS 2280

Db 2221 MFSISGRRTMIHPGVNRSSSTSPVSKGKPLKTPASKSPSEGTATTSRPGAKPSVKS 2280

Qy 2281 ELSPVARTQISGSGSKAPSRSGSDSTPSRPAQPLSRPIQSPGRNLSIFGRNGISPPN 2340

Db 2281 ELSPVARTQISGSGSKAPSRSGSDSTPSRPAQPLSRPIQSPGRNLSIFGRNGISPPN 2340

Qy 2341 KLSQLPRTSSPTASTKSGSGSKMYSITSPGRMSQNLTKOTGLSKNASSIPRESASKG 2400

Db 2341 KLSQLPRTSSPTASTKSGSGSKMYSITSPGRMSQNLTKOTGLSKNASSIPRESASKG 2400

Qy 2401 LNMNNGANKKVELSRMSTKSGSGSDRSERPVLRQSTTFKEAPSPTLRRKLEESA 2460

Db 2401 LNMNNGANKKVELSRMSTKSGSGSDRSERPVLRQSTTFKEAPSPTLRRKLEESA 2460

Qy 2461 SPESLSPSRPASPTRSQAQTPVLSPLPDYSLSTHSSVQAGGWRKLPNLSPTIENDG 2520

Db 2461 SPESLSPSRPASPTRSQAQTPVLSPLPDYSLSTHSSVQAGGWRKLPNLSPTIENDG 2520

Qy 2521 RPAKRHDIAHSHSESPRLPINRSGTWKREHSKSSSLPRVSTWRRRTGSSSSILSASSES 2580

Db 2521 RPAKRHDIAHSHSESPRLPINRSGTWKREHSKSSSLPRVSTWRRRTGSSSSILSASSES 2580

Qy 2581 SEKAKSEBKHVNSLSGTQSKENQVSAKGTWRKIKENEFPTNSTQTSVSSGATNGAES 2640

Db 2581 SEKAKSEBKHVNSLSGTQSKENQVSAKGTWRKIKENEFPTNSTQTSVSSGATNGAES 2640

Qy 2641 KTLIIQMAPAVSKTIDVWYRIEDCPINNRSGRSPGTNTPPVIDSVSEKAPNLIKSDKN 2700

Db 2641 KTLIIQMAPAVSKTIDVWYRIEDCPINNRSGRSPGTNTPPVIDSVSEKAPNLIKSDKN 2700

Qy 2701 QAKQNVGSGVPMRTVGLNRLNSFTQVDAPQKGTETIKPGQNNPVPVSETNESSIVERT 2760

Db 2701 QAKQNVGSGVPMRTVGLNRLNSFTQVDAPQKGTETIKPGQNNPVPVSETNESSIVERT 2760

Qy 2761 PFSSSSSKHSPSGTVAARVTPFNPNPGRKSSADSTARPSPQIPTPVNNNTKRDST 2820

121 RGVNGSRSTGYLEELKERSLILLADLDKEEKEDWYVAQLONLTKRIDSLPTENFSL 180  
 121 RGVNGSRSTGYLEELKERSLILLADLDKEEKEDWYVAQLONLTKRIDSL- LTIENFSL 179  
 181 QTMTRQLRYEARQIRVAMEBOLGTCQDMKEKRAQRIARIQOIEKDIIRIROLLOSQT 240  
 180 QTMTRQLRYEARQIRVAMEBOLGTCQDMKEKRAQRIARIQOIEKDIIRIROLLOSQT 239  
 241 EARSSONKHETGSHDAERQOGVEINMATSGNGQSGSTTRMDHETASVLSSSSTHSA 300  
 240 EARSSONKHETGSHDAERQOGVEINMATSGNGQSGSTTRMDHETASVLSSSSTHSA 299  
 301 PRRLTSHLGTKEVNVYSLLSMLGTHDKDMSRLLAMSSQDSQSCINRQSGCLPLLIQLL 360  
 300 PRRLTSHLGTKEVNVYSLLSMLGTHDKDMSRLLAMSSQDSQSCINRQSGCLPLLIQLL 359  
 361 HGNDKDSVLGNSRGSKEARASAAALHNIHSQDDKGRREIRVLHLEQIRAYCETC 420  
 360 HGNDKDSVLGNSRGSKEARASAAALHNIHSQDDKGRREIRVLHLEQIRAYCETC 419  
 421 WEMQEAHEPGWDQKNPMPAPVEHQICPAVCVLMKLSFDEEHRHAMNELGGLQAI AELLQ 480  
 420 WEMQEAHEPGWDQKNPMPAPVEHQICPAVCVLMKLSFDEEHRHAMNELGGLQAI AELLQ 479  
 481 VDCMYGLTNDHYSITLRRYAGHALTNLTGPDVANKATILCSMGCMRALVAQLKSEBDL 540  
 480 VDCMYGLTNDHYSITLRRYAGHALTNLTGPDVANKATILCSMGCMRALVAQLKSEBDL 539  
 541 QOVLASVLRNLWRADVNSKKTREVGSVKALMECALEVKKESTLXSVLSALWNLSAHCT 600  
 540 QOVLASVLRNLWRADVNSKKTREVGSVKALMECALEVKKESTLXSVLSALWNLSAHCT 599  
 601 ENKADI CAVDGALAFVLGTYRSTQNTLAIIESGGGILNVSSLIATNEDHROILRENN 660  
 600 ENKADI CAVDGALAFVLGTYRSTQNTLAIIESGGGILNVSSLIATNEDHROILRENN 659  
 661 CLOTLLOHLKSHSLTIVSNACGLTNLNSARNPKDQALWDMGAVSMKMLIHSKHMIAM 720  
 660 CLOTLLOHLKSHSLTIVSNACGLTNLNSARNPKDQALWDMGAVSMKMLIHSKHMIAM 719  
 721 GSAALRNLMANRPKADANINSPGSLPSLHVRKQKALEAELDAQHLSSETFDNIDNLS 780  
 720 GSAALRNLMANRPKADANINSPGSLPSLHVRKQKALEAELDAQHLSSETFDNIDNLS 779  
 781 PKASHRSKQHKOSLYGDIYFDNRHDDNDSNFTGNMTVLSPLYLNTTVLPSSSSSRGS 840  
 780 PKASHRSKQHKOSLYGDIYFDNRHDDNDSNFTGNMTVLSPLYLNTTVLPSSSSSRGS 839  
 841 LDSRSSEKSLRERERIGLGNVHPATENPGTSSKRGLOISTTAAQIAKVMEEVSAIHTS 900  
 840 LDSRSSEKSLRERERIGLGNVHPATENPGTSSKRGLOISTTAAQIAKVMEEVSAIHTS 899  
 901 QEDRSSGTTLHCVTERNALRSSAAHSTNYFTKSENSNRCTCSMPYAKLEYKRS 960  
 900 QEDRSSGTTLHCVTERNALRSSAAHSTNYFTKSENSNRCTCSMPYAKLEYKRS 959  
 961 NDSLNSVSSDGYKRGOMKPSIESYEDDESFCVGYQPADLAHKIHSANHMDNDGE 1020  
 960 NDSLNSVSSDGYKRGOMKPSIESYEDDESFCVGYQPADLAHKIHSANHMDNDGE 1019  
 1021 LDTPINYSKYDBQLNSGRQSPQNERWARPKHIIIEDEIKQSEQRNQSTTYPVYTE 1080  
 1020 LDTPINYSKYDBQLNSGRQSPQNERWARPKHIIIEDEIKQSEQRNQSTTYPVYTE 1079  
 1081 STDDKHLKFOHFGQOCVFPYRSRGANGSETNRVGSNHGINVQSLOCOEDDYEDDKP 1140  
 1080 STDDKHLKFOHFGQOCVFPYRSRGANGSETNRVGSNHGINVQSLOCOEDDYEDDKP 1139  
 1141 TNSYERYSEEQHEEERPTNYSIKYNEEKRVDPIDYSLKYATDIPSSQKQSFSSKS 1200  
 1140 TNSYERYSEEQHEEERPTNYSIKYNEEKRVDPIDYSLKYATDIPSSQKQSFSSKS 1199  
 1201 SSCOSSKTHWSSSFTWSTPSSNAKPNOLHPSSASPSGQOKAATCKVSSINQETIQ 1260

1200 SSGQSSKTHWSSSFTWSTPSSNAKPNOLHPSSASPSGQOKAATCKVSSINQETIQ 1259  
 1261 TYCYVEDTPICFSRCSLSLSLSAEDEIGCNOTTQOADSANTLQIAEIKKIGTRSAADPV 1320  
 1260 TYCYVEDTPICFSRCSLSLSLSAEDEIGCNOTTQOADSANTLQIAEIKKIGTRSAADPV 1319  
 1321 SEVPAVSQHPRTKSRSLQSSLSSESARHKAVERFSSGAKSPSKSGAQTTPSPPHYVOET 1380  
 1320 SEVPAVSQHPRTKSRSLQSSLSSESARHKAVERFSSGAKSPSKSGAQTTPSPPHYVOET 1379  
 1381 PLMFSRCTSVSSLSDFSRSTASSVQSEPCSGMVGIIISPDLPDSPGQTMPPSRSTPP 1440  
 1380 PLMFSRCTSVSSLSDFSRSTASSVQSEPCSGMVGIIISPDLPDSPGQTMPPSRSTPP 1439  
 1441 PPPOTACTKREVPKXKAPTAEKRESGPKQAANVAQVQVLPDADTLHLHFATESTPDGF 1500  
 1440 PPPOTACTKREVPKXKAPTAEKRESGPKQAANVAQVQVLPDADTLHLHFATESTPDGF 1499  
 1501 SCSSLSLSLDEPFIQKVELRIMPPVOENDNGNETESQPKESNENQEKBAKTTIDSE 1560  
 1500 SCSSLSLSLDEPFIQKVELRIMPPVOENDNGNETESQPKESNENQEKBAKTTIDSE 1559  
 1561 KOLLDDSDDDIIELEECISAMPTKSRKAKKPAQATASKLPPPVARKPSQOLPVYKLLPS 1620  
 1560 KOLLDDSDDDIIELEECISAMPTKSRKAKKPAQATASKLPPPVARKPSQOLPVYKLLPS 1619  
 1621 QNRLOPKHVSFTPGDDMPRVYCVVEGTPINPSTATSLDLTIESPPNELAAGEVGRGAQ 1680  
 1620 QNRLOPKHVSFTPGDDMPRVYCVVEGTPINPSTATSLDLTIESPPNELAAGEVGRGAQ 1679  
 1681 SGEFEKRTITIEGSRSTDEAOGGKTSVTIPELDDNKAEGDILAEICINSAMPKKGSHKP 1740  
 1680 SGEFEKRTITIEGSRSTDEAOGGKTSVTIPELDDNKAEGDILAEICINSAMPKKGSHKP 1739  
 1741 FRVKXIMDQVQOASASSAPKNQLDGKKKPTSPVKPIPQNTFYRTRVRKNADSKNKLN 1800  
 1740 FRVKXIMDQVQOASASSAPKNQLDGKKKPTSPVKPIPQNTFYRTRVRKNADSKNKLN 1799  
 1801 AERVFSDNKSQKQNLKNNKDFNDKLPNNEDRVGSGFAEDSPHHYTPIEGTPYCFSRND 1860  
 1800 AERVFSDNKSQKQNLKNNKDFNDKLPNNEDRVGSGFAEDSPHHYTPIEGTPYCFSRND 1859  
 1861 SLSLDFDDDDVDLSREKAELRKAKENKESAKVTSHTELTSNQOSANKTQAIKQPINR 1920  
 1860 SLSLDFDDDDVDLSREKAELRKAKENKESAKVTSHTELTSNQOSANKTQAIKQPINR 1919  
 1921 GQPKPILOKQTFPQSSKDIPDRGAATDEKLQNFATIENTPVCFSHNSSLSSLSLSDIDENN 1980  
 1920 GQPKPILOKQTFPQSSKDIPDRGAATDEKLQNFATIENTPVCFSHNSSLSSLSLSDIDENN 1979  
 1981 NKENEPIKETEPPOSQOGEPSKQAGVAPKSFHVEDTPVCFSRNSSLSLSDIDEDDLQ 2040  
 1980 NKENEPIKETEPPOSQOGEPSKQAGVAPKSFHVEDTPVCFSRNSSLSLSDIDEDDLQ 2039  
 2041 ECISAMPKX 2100  
 2040 ECISAMPKX 2099  
 2101 WKATOEGANSIVSSIHQAAAAACLSRQASDSDSILSKSGISLGSFPFLTTPDOEEKPFT 2160  
 2100 WKATOEGANSIVSSIHQAAAAACLSRQASDSDSILSKSGISLGSFPFLTTPDOEEKPFT 2159  
 2161 SNKGPRILKPEKSTLETTKIETESKGIKGGKVKYKSLITGKVESNSEIISGOMKQPIQAN 2220  
 2160 SNKGPRILKPEKSTLETTKIETESKGIKGGKVKYKSLITGKVESNSEIISGOMKQPIQAN 2219  
 2221 MPSISRGTHIHIHGVANSSSTSPVSKKGPPLTKPASKSPSEGTATTTPRGAKPVKS 2280  
 2220 MPSISRGTHIHIHGVANSSSTSPVSKKGPPLTKPASKSPSEGTATTTPRGAKPVKS 2279  
 2281 ELSPVARTQSIQAGSSKAPSRSGSRDTPSPAPQPSRPIQSPGRNISIPGRNGISPPN 2340

Db 2280 ELSPVARTSQIGSSGKAPSRSGSRDSTPSRPAQQLRPIQSPGRNSISPRNGISPPN 2339  
Qy 2341 KLSQLPRTSSPSTASTKSGGKGMQYTSFGQMSQONLTQTGLSKNASSIPRSESAGK 2400  
Db 2340 KLSQLPRTSSPSTASTKSGGKGMQYTSFGQMSQONLTQTGLSKNASSIPRSESAGK 2399  
Qy 2401 LNOMNGNGANKVLSRMSSTKSSGSESDRSERPVLVRQSTFFIKEAPSPTLRKLBSA 2460  
Db 2400 LNOMNGNGANKVLSRMSSTKSSGSESDRSERPVLVRQSTFFIKEAPSPTLRKLBSA 2459  
Qy 2461 SPESLSPSRPASPRSAQOTPVLSPLPDMSLTHSSVOAGGWRKLPNLSPTIEYNDG 2520  
Db 2460 SPESLSPSRPASPRSAQOTPVLSPLPDMSLTHSSVOAGGWRKLPNLSPTIEYNDG 2519  
Qy 2521 RPAKHHDARSSESPLPINSRGTWKRHSKHSSSLPRVSTWRTTSGSSSIISSASES 2580  
Db 2520 RPAKHHDARSSESPLPINSRGTWKRHSKHSSSLPRVSTWRTTSGSSSIISSASES 2579  
Qy 2581 SEKASEDEKHVNSISGTSKQSKENQVSAKGTWRKIKENEFPTNSTQTSVSGATNGAES 2640  
Db 2580 SEKASEDEKHVNSISGTSKQSKENQVSAKGTWRKIKENEFPTNSTQTSVSGATNGAES 2639  
Qy 2641 KTLIYOMAPAVSKTDVWRIEDCPINNPRSGRSPGTNTPPVIDSVSEKANPNIKOSKDN 2700  
Db 2640 KTLIYOMAPAVSKTDVWRIEDCPINNPRSGRSPGTNTPPVIDSVSEKANPNIKOSKDN 2699  
Qy 2701 QAKQNVGNGSVPMRTVGLNRLNSFIQVADPQKGTETKPGONNPVPVSETNESSIVERT 2760  
Db 2700 QAKQNVGNGSVPMRTVGLNRLNSFIQVADPQKGTETKPGONNPVPVSETNESSIVERT 2759  
Qy 2761 PFSSSSSKHSHSPSGTVAARVTPFNPNPSRKSADSTSAEPSQIPTPVNNNTKKRDSKT 2820  
Db 2760 PFSSSSSKHSHSPSGTVAARVTPFNPNPSRKSADSTSAEPSQIPTPVNNNTKKRDSKT 2819  
Qy 2821 DSTESSGTQSPKRHSGSVLTVSV 2843  
Db 2820 DSTESSGTQSPKRHSGSVLTVSV 2842

RESULT 9

US-08-289-548A-7

Sequence 7, Application US/08289546A

Patent No. 5648212

GENERAL INFORMATION:

APPLICANT: ALBERTSEN, HANS

APPLICANT: ANAND, RAKESH

APPLICANT: CARLSON, MARY

APPLICANT: GRODEN, JOANNA

APPLICANT: HEDGE, PHILIP J.

APPLICANT: JOSLYN, GEORF

APPLICANT: KINZLER, KENNETH

APPLICANT: MARKHAM, ALEXANDER F.

APPLICANT: NAKAMURA, YUSUKE

APPLICANT: THLIVERIS, ANDREW

TITLE OF INVENTION: INHERITED AND SOMATIC MUTATIONS OF APC

TITLE OF INVENTION: GENE IN COLORECTAL CANCER IN HUMANS

NUMBER OF SEQUENCES: 102

CORRESPONDENCE ADDRESS:

ADDRESSEE: Banner & Allegretti, LTD

STREET: 1001 G Street, NW

CITY: Washington

STATE: D.C.

COUNTRY: USA

ZIP: 20001-4598

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/289,548A

FILING DATE: 12-AUG-1994

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Kagan, Sarah A.

REGISTRATION NUMBER: 32,141

REFERENCE/DOCKET NUMBER: 1107.46943

TELECOMMUNICATION INFORMATION:

TELEPHONE: 202-508-9100

TELEFAX: 202-508-9299

INFORMATION FOR SEQ ID NO: 7:

SEQUENCE CHARACTERISTICS:

LENGTH: 2842 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

ORIGINAL SOURCE:

ORGANISM: Homo sapiens

IMMEDIATE SOURCE:

CLONE: APC

US-08-289-548A-7

Query Match 99.9%; Score 14548.5; DB 1; Length 2842;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 2842; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

Qy 1 MAAASVDQLLKQVEALKMENSNIROBLEDNSNHLTKLETEASNKKEVLKQLOGSIEDEAM 60

Db 1 MAAASVDQLLKQVEALKMENSNIROBLEDNSNHLTKLETEASNKKEVLKQLOGSIEDEAM 60

Qy 61 ASSQIDLLERLKLNLDSNPFQVKLRKWSLRSYGSRGVSYSRSGECSVPYMGSPFR 120

Db 61 ASSQIDLLERLKLNLDSNPFQVKLRKWSLRSYGSRGVSYSRSGECSVPYMGSPFR 120

Qy 121 RGVNRSRESTGYLEELERKERSLLADLDKEEKEKDWYAAQLQNLTKRIDSLPTNFSL 180

Db 121 RGVNRSRESTGYLEELERKERSLLADLDKEEKEKDWYAAQLQNLTKRIDSLPTNFSL 180

Qy 181 QDTMTRQLEYEARQIRVAMEEQLGTCCODMEKAQRIARIQOIEKDILRIQLQSOAT 240

Db 180 QDTMTRQLEYEARQIRVAMEEQLGTCCODMEKAQRIARIQOIEKDILRIQLQSOAT 239

Qy 241 EAERSSQNHETGSHDAERQNEGQVGEINWATSGNGQSTTMDHETASVLSSSSTHSA 300

Db 240 EAERSSQNHETGSHDAERQNEGQVGEINWATSGNGQSTTMDHETASVLSSSSTHSA 299

Qy 301 PRLTSHLGTKEVMYVSYLLSMLGTHDKDMSRTLLAMSSQSDSCISMRQGCPLLIQLL 360

Db 300 PRLTSHLGTKEVMYVSYLLSMLGTHDKDMSRTLLAMSSQSDSCISMRQGCPLLIQLL 359

Qy 361 HGNDKDSVLLGNSRGSKERARASAAALHNIHSQPDQKGRREIRVHLLEQIRAYCETC 420

Db 360 HGNDKDSVLLGNSRGSKERARASAAALHNIHSQPDQKGRREIRVHLLEQIRAYCETC 419

Qy 421 WEMQEAHEPGMDQKMPMPAPVEHQICPAVCVLMKLSFDEEHRHAMNELGGLQAIALLQ 480

Db 420 WEMQEAHEPGMDQKMPMPAPVEHQICPAVCVLMKLSFDEEHRHAMNELGGLQAIALLQ 479

Qy 481 VDCMYGLTNDHYSITLRRYAGMALTNLTGPDVANKATLCMKGCMRALVAQKSEEDL 540

Db 480 VDCMYGLTNDHYSITLRRYAGMALTNLTGPDVANKATLCMKGCMRALVAQKSEEDL 539

Qy 541 QQVIASVLRNLSWRADVNSKKTIREVGSVKALMECALEVKKESTLKSLSALWNLSAHCT 600

Db 540 QQVIASVLRNLSWRADVNSKKTIREVGSVKALMECALEVKKESTLKSLSALWNLSAHCT 599

Qy 601 ENKADICAVDGAFLVGLTYRSQNTNTLAIIESGGGILRNVSLLATNEDHQIILRENN 660

Db 600 ENKADICAVDGAFLVGLTYRSQNTNTLAIIESGGGILRNVSLLATNEDHQIILRENN 659

Qy 661 CLQTLLOHLKSHSLTIIVSNACGLTNLSARNPKDQBALWDMGVSMKLNLIHSHKHMAM 720

Db 660 CLQTLLOHLKSHSLTIIVSNACGLTNLSARNPKDQBALWDMGVSMKLNLIHSHKHMAM 719

Qy 721 GSAAALENIMANRPKDYKDNIMSPGSSLSLHVVRKQKALEAEALDAHLSETFDNIDNLS 780

Db 720 GSAALNLANRPAKYKDIANIMSPGSSLSLVRKQKALEADLQHLSETFDNIDNLS 779  
Qy 781 PKASHRSKQKHQSLGYDYVFTNDRHDDNRSDFNTGNMTVLSPYLNTTVLPSSSSSRGS 840  
Db 780 PKASHRSKQKHQSLGYDYVFTNDRHDDNRSDFNTGNMTVLSPYLNTTVLPSSSSSRGS 839  
Qy 841 LDSRSKSLRLERIGIGLGNTHPATENPGTSSKRGLOLSTTAAQIAKMEVSAIHTS 900  
Db 840 LDSRSKSLRLERIGIGLGNTHPATENPGTSSKRGLOLSTTAAQIAKMEVSAIHTS 899  
Qy 901 QEDRSSGTTTELHCVTDERNALRRSSAAHSTNYNFTKSENRRCTSMPIAKLEYKRSS 960  
Db 900 QEDRSSGTTTELHCVTDERNALRRSSAAHSTNYNFTKSENRRCTSMPIAKLEYKRSS 959  
Qy 961 NDSLNSVSSDGYGKRGOMKPSIESYSEDDESKFCYGYQYPADLAHKIHSANHMDNDGE 1020  
Db 960 NDSLNSVSSDGYGKRGOMKPSIESYSEDDESKFCYGYQYPADLAHKIHSANHMDNDGE 1019  
Qy 1021 LDTPINYLKYSDBOLNSGRQSPSONERWARPKHIIIEDEIKQSEQRNQSTTYPVYTE 1080  
Db 1020 LDTPINYLKYSDBOLNSGRQSPSONERWARPKHIIIEDEIKQSEQRNQSTTYPVYTE 1079  
Qy 1081 STDDKHLKFQPHFQOCVCVPSYRSGANGSETNRVGSNHNQINQVQSLOEDEDYEDDKP 1140  
Db 1080 STDDKHLKFQPHFQOCVCVPSYRSGANGSETNRVGSNHNQINQVQSLOEDEDYEDDKP 1139  
Qy 1141 TNSYERYSEBQHEBERPTNYIKNEEKRHVDQPIDYSLKYATDIPSOQKQSPFSKS 1200  
Db 1140 TNSYERYSEBQHEBERPTNYIKNEEKRHVDQPIDYSLKYATDIPSOQKQSPFSKS 1199  
Qy 1201 SSGQSKTEHWSSENSTPSSNAKQONLHPSSAQSRGQPKAATCKVSSINQETIQ 1260  
Db 1200 SSGQSKTEHWSSENSTPSSNAKQONLHPSSAQSRGQPKAATCKVSSINQETIQ 1259  
Qy 1261 TYCVEDTPICFRCSLSLSSAEDIEGICNQTQOEAASANTLOIAIEIKKIGITRSAEDPV 1320  
Db 1260 TYCVEDTPICFRCSLSLSSAEDIEGICNQTQOEAASANTLOIAIEIKKIGITRSAEDPV 1319  
Qy 1321 SEVPAVSQHRTKSSRLQGSLSLSESAHKAIVFSSGAKSPSKSGAOTPKSPPEHVQET 1380  
Db 1320 SEVPAVSQHRTKSSRLQGSLSLSESAHKAIVFSSGAKSPSKSGAOTPKSPPEHVQET 1379  
Qy 1381 PLMFSTRCTSVSSLDSPESRSIASSVQSEPCSGMVSGIIPSDDLDPSPGOTMPPSRKTPP 1440  
Db 1380 PLMFSTRCTSVSSLDSPESRSIASSVQSEPCSGMVSGIIPSDDLDPSPGOTMPPSRKTPP 1439  
Qy 1441 PPQTAQTKREVPKNKAAPTAKRESGPKQAANVAQVQVLPDADTLHLFAESTPDGF 1500  
Db 1440 PPQTAQTKREVPKNKAAPTAKRESGPKQAANVAQVQVLPDADTLHLFAESTPDGF 1499  
Qy 1501 SCSSLSALSILDEPFIQKVELRIMPVQENDNGNETESEQKESNENQKAEKTIIDSE 1560  
Db 1500 SCSSLSALSILDEPFIQKVELRIMPVQENDNGNETESEQKESNENQKAEKTIIDSE 1559  
Qy 1561 KDLLDDSDDDDDIELEECIIISAMPTKSSRKAKPAQATASKLPPPVARKPSQLPVYKLLPS 1620  
Db 1560 KDLLDDSDDDDDIELEECIIISAMPTKSSRKAKPAQATASKLPPPVARKPSQLPVYKLLPS 1619  
Qy 1621 QNRLOPKHVSFTPGDDMPRVYCVEGTPIINFSTATSLSDLTIESPPNELAAGEGVGGGAQ 1680  
Db 1620 QNRLOPKHVSFTPGDDMPRVYCVEGTPIINFSTATSLSDLTIESPPNELAAGEGVGGGAQ 1679  
Qy 1681 SGBFEKEDTITTEGRSTDEAQGKTSVITPELDDNKAEEDGILABEINCSAMPKGSHPK 1740  
Db 1680 SGBFEKEDTITTEGRSTDEAQGKTSVITPELDDNKAEEDGILABEINCSAMPKGSHPK 1739  
Qy 1741 FRVKKINDVQOASASSAPNKNQLODKKKKTPSPVKEPIQNTYEYTRVRKNADSKNNLN 1800  
Db 1740 FRVKKINDVQOASASSAPNKNQLODKKKKTPSPVKEPIQNTYEYTRVRKNADSKNNLN 1799  
Qy 1801 AERVFSNKKDSKKNLKNNSKDFNDKLPNNEDVRGSAFDPSPHYTPIEGTPYCFSRND 1860

Db 1800 AERVFSNKKDSKKNLKNNSKDFNDKLPNNEDVRGSAFDPSPHYTPIEGTPYCFSRND 1859  
Qy 1861 SILSLDFDDDDVDLSRKAELRKAKENKESAKVTSHTELTSNQOSANKTOIAIAKQPINR 1920  
Db 1860 SILSLDFDDDDVDLSRKAELRKAKENKESAKVTSHTELTSNQOSANKTOIAIAKQPINR 1919  
Qy 1921 COPRPILOKQSTFPQSSKDIIPDRGAATDEKLOFAIENTPVCFSHNSLSLSLSDIDQENN 1980  
Db 1920 COPRPILOKQSTFPQSSKDIIPDRGAATDEKLOFAIENTPVCFSHNSLSLSLSDIDQENN 1979  
Qy 1981 NKENEPKETETPPDOSQGBEPSPQASGVAPKSFHVEDTPVCFSRNSSLSLSLSDIDEDDILQ 2040  
Db 1980 NKENEPKETETPPDOSQGBEPSPQASGVAPKSFHVEDTPVCFSRNSSLSLSLSDIDEDDILQ 2039  
Qy 2041 ECISAMPKPKKPRLLKGDNEKHSFRMNGGILGDEDLTLDKDIQRPDSEHGLSPDSENF 2100  
Db 2040 ECISAMPKPKKPRLLKGDNEKHSFRMNGGILGDEDLTLDKDIQRPDSEHGLSPDSENF 2099  
Qy 2101 WKAIQEGANSIVSSLHOAAAAACLSROASDSDSILSKSGISLGSPPHLPDQOEKPPFT 2160  
Db 2100 WKAIQEGANSIVSSLHOAAAAACLSROASDSDSILSKSGISLGSPPHLPDQOEKPPFT 2159  
Qy 2161 SNKGPRILKPGCKSTLETKKIESESKGKGGKYYKSLITGKVRNSNBEISGQMKQPLQAN 2220  
Db 2160 SNKGPRILKPGCKSTLETKKIESESKGKGGKYYKSLITGKVRNSNBEISGQMKQPLQAN 2219  
Qy 2221 MPSISRGRTMIHPGVNRNSSSTSPVSKKGPPLTPASKSPSEGTATTPRGAKPSPVKS 2280  
Db 2220 MPSISRGRTMIHPGVNRNSSSTSPVSKKGPPLTPASKSPSEGTATTPRGAKPSPVKS 2279  
Qy 2281 ELSPVARTSIOGSSKAPSRSGSDSTPSRPAQOPLSRPIQSPGRNSISPGRNGISPPN 2340  
Db 2280 ELSPVARTSIOGSSKAPSRSGSDSTPSRPAQOPLSRPIQSPGRNSISPGRNGISPPN 2339  
Qy 2341 KLSQLPRTSPSTASTKSSGSGKMSYTSPOGMSQOQLTKQTGLSKNASSIPRSESASK 2400  
Db 2340 KLSQLPRTSPSTASTKSSGSGKMSYTSPOGMSQOQLTKQTGLSKNASSIPRSESASK 2399  
Qy 2401 LNQMNGANGANKVELSRMSTKSSGESRDERPVLVRQSTFIKEAPSTLRRKLBSA 2460  
Db 2400 LNQMNGANGANKVELSRMSTKSSGESRDERPVLVRQSTFIKEAPSTLRRKLBSA 2459  
Qy 2461 SFESLSPSRSPASPTRSQOATPVLPSLPMGLSTHSSVQAGWRKLPNLSPTIEYNDG 2520  
Db 2460 SFESLSPSRSPASPTRSQOATPVLPSLPMGLSTHSSVQAGWRKLPNLSPTIEYNDG 2519  
Qy 2521 RPAKHDIARSHSPSRPLPNRSGTWKRHSKSSSLPRVSTWRRRTGSSSSILSASSES 2580  
Db 2520 RPAKHDIARSHSPSRPLPNRSGTWKRHSKSSSLPRVSTWRRRTGSSSSILSASSES 2579  
Qy 2581 SEKAKSEDEKHVNSISGTQSKENQVSAKGTWRKIKENEFSPNTNSTQTVSSGATNGAES 2640  
Db 2580 SEKAKSEDEKHVNSISGTQSKENQVSAKGTWRKIKENEFSPNTNSTQTVSSGATNGAES 2639  
Qy 2641 KTLIYQMAPAVSKTEDVWVRIEDCPINNPRSGRSPCTNTPTVIDSVSEKANPNIKSKDN 2700  
Db 2640 KTLIYQMAPAVSKTEDVWVRIEDCPINNPRSGRSPCTNTPTVIDSVSEKANPNIKSKDN 2699  
Qy 2701 QAKQNVNGSVPMRTVGLNRLNSFIOVDAPDQKGTETIKPGQNNPVPVSETNESSIVERT 2760  
Db 2700 QAKQNVNGSVPMRTVGLNRLNSFIOVDAPDQKGTETIKPGQNNPVPVSETNESSIVERT 2759  
Qy 2761 PFSSSSSKSHSPSGTVAARVTPPNYNPSPRKSSADTSARPQOIPTPVNNNTKKRDSKT 2820  
Db 2760 PFSSSSSKSHSPSGTVAARVTPPNYNPSPRKSSADTSARPQOIPTPVNNNTKKRDSKT 2819  
Qy 2821 DSTESSGTQSPKHSRSGSLVTSV 2843  
Db 2820 DSTESSGTQSPKHSRSGSLVTSV 2842

```

; Sequence 7, Application US/08452654
; Patent No. 5691454
; GENERAL INFORMATION:
; APPLICANT: ALBERTSEN, HANS
; APPLICANT: ANAND, RAKESH
; APPLICANT: CARLSON, MARY
; APPLICANT: GRODEN, JOANNA
; APPLICANT: HEDGE, PHILIP J.
; APPLICANT: JOSLYN, GEOFF
; APPLICANT: KINZLER, KENNETH
; APPLICANT: NAKHAM, ALEXANDER F.
; APPLICANT: NAKAMURA, YUSUKE
; APPLICANT: THLIVERTIS, ANDREW
; TITLE OF INVENTION: INHERITED AND SOMATIC MUTATIONS OF APC
; TITLE OF INVENTION: GENE IN COLORECTAL CANCER IN HUMANS
; NUMBER OF SEQUENCES: 94
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Banner, Birch, McKie & Beckett
; STREET: 1001 G Street, NW
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20001-4598
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/452,654
; FILING DATE: 25-MAY-1995
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/741,940
; FILING DATE: 08-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Kagan, Sarah A.
; REGISTRATION NUMBER: 32,141
; REFERENCE/DOCKET NUMBER: 1107.035574
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-508-9100
; TELEFAX: 202-508-9299
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2842 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; IMMEDIATE SOURCE:
; CLONE: APC
; US-08-452-654-7

Query Match 99.9%; Score 14548.5; DB 1; Length 2842;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2842; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 MAASVDQLLKVEALKMNSNLRQLEDNSNHLTKLEASNMKEVLKQLQGSSTEDAM 60
DB 1 MAASVDQLLKVEALKMNSNLRQLEDNSNHLTKLEASNMKEVLKQLQGSSTEDAM 60
QY 61 ASSGQIDLLERLKEALNDSSNPFPGVKLRSMKSLRSGREGSVSSRSGECSFVPMGSPFR 120
DB 61 ASSGQIDLLERLKEALNDSSNPFPGVKLRSMKSLRSGREGSVSSRSGECSFVPMGSPFR 120
QY 121 RGFVNGSRESTGYLBELEKERSILLADLDKEKEKDWYQAQLNLTKEIDSLPTENESL 180
DB 121 RGFVNGSRESTGYLBELEKERSILLADLDKEKEKDWYQAQLNLTKEIDSLPTENESL 179
QY 181 QDTMTRRQLEYEARQIRVAMEEQLGTCQDMKEKAAQRRIARIQOIEKDIILRIQLLSQAT 240

```

```

180 QDTMTRRQLEYEARQIRVAMEEQLGTCQDMKEKAAQRRIARIQOIEKDIILRIQLLSQAT 239
241 EAERSQNKHETGSHDAERQNEGGVGEINMATSGNGQSGTTRMDHETASVLSSSTHSA 300
240 EAERSQNKHETGSHDAERQNEGGVGEINMATSGNGQSGTTRMDHETASVLSSSTHSA 299
301 PRRLTSHLGTKEVMYVSLLSMLGTHDKDDYSRTLLAMSSQDSCISMRSGCCLPLLIQLL 360
300 PRRLTSHLGTKEVMYVSLLSMLGTHDKDDYSRTLLAMSSQDSCISMRSGCCLPLLIQLL 359
361 HGNDKDSVLLGNRSRGSKEARASAAALHNIHSQPDQDKRGRREIRVJLHLEQIRAYCETC 420
360 HGNDKDSVLLGNRSRGSKEARASAAALHNIHSQPDQDKRGRREIRVJLHLEQIRAYCETC 419
421 NEWQEAHEPGMDQKMPMPAPVEHQICPAVCULMKLSFDEEHRHMANVELGGQAIABELQ 480
420 NEWQEAHEPGMDQKMPMPAPVEHQICPAVCULMKLSFDEEHRHMANVELGGQAIABELQ 479
481 VDCEMYGLTNDHYSITLRRYAGMALTNLTFDGVANKATLCSMKGCMRALVAQLKSSSEDL 540
480 VDCEMYGLTNDHYSITLRRYAGMALTNLTFDGVANKATLCSMKGCMRALVAQLKSSSEDL 539
541 QQVIASVLRNLSWRADVNSKTLREVGSKVAMECALEVKKESTLKSIVLSALWNLSAHCT 600
540 QQVIASVLRNLSWRADVNSKTLREVGSKVAMECALEVKKESTLKSIVLSALWNLSAHCT 599
601 ENKADICAVDGAFLVGLTYESQNTNLAIIESGGILRNVSLLIATNEDHRCIILRENN 660
600 ENKADICAVDGAFLVGLTYESQNTNLAIIESGGILRNVSLLIATNEDHRCIILRENN 659
661 CLQTLQLKLSHSLTIVSNACGLTNLSARNPKDQBALMDMGAVSMLKNLIHSHKHMIAM 720
660 CLQTLQLKLSHSLTIVSNACGLTNLSARNPKDQBALMDMGAVSMLKNLIHSHKHMIAM 719
721 GSAAALRNLMANRPKAKYKDANINSPGSSLPVLHVRKOKALEAELDAHLSEFNDNLIS 780
720 GSAAALRNLMANRPKAKYKDANINSPGSSLPVLHVRKOKALEAELDAHLSEFNDNLIS 779
781 PKASHRSKQKHQKSLYGVYFDNRRHDDNRSDNFNTGNMTVLSPLYNTTLVLPSSSSSRGS 840
780 PKASHRSKQKHQKSLYGVYFDNRRHDDNRSDNFNTGNMTVLSPLYNTTLVLPSSSSSRGS 839
841 LDSSRSEKORSLERGIGLGNTHPATENPGTSKKGLOISTTAAQIAKMEVSAIHIS 900
840 LDSSRSEKORSLERGIGLGNTHPATENPGTSKKGLOISTTAAQIAKMEVSAIHIS 899
901 QEDRSSGSTTELHCVTDERNALRSSAAHSTNFTKSENSNRCTCSFYAKLEYKRSS 960
900 QEDRSSGSTTELHCVTDERNALRSSAAHSTNFTKSENSNRCTCSFYAKLEYKRSS 959
961 NDSLNSVSSSDGYGKRGQMKPSIESYSEDESKFCYGYPADLAHKIHSANHMDNDGE 1020
960 NDSLNSVSSSDGYGKRGQMKPSIESYSEDESKFCYGYPADLAHKIHSANHMDNDGE 1019
1021 LDTPINYSKYDEOLNSGRQSPSONERWAPKHIIIEDEIKOSEQORSNQSTTYVYTE 1080
1020 LDTPINYSKYDEOLNSGRQSPSONERWAPKHIIIEDEIKOSEQORSNQSTTYVYTE 1079
1081 STDDKHLKFQPHFGQCEVSPYASRGANGSETNRVNGSHNGINQVNSQSLCQEDDDYDDKP 1140
1080 STDDKHLKFQPHFGQCEVSPYASRGANGSETNRVNGSHNGINQVNSQSLCQEDDDYDDKP 1139
1141 TNYSERYSBEOQHEBERPTNYSIKYNEEKHVDQIDYSLKYATDIPSSQKGSFSPKS 1200
1140 TNYSERYSBEOQHEBERPTNYSIKYNEEKHVDQIDYSLKYATDIPSSQKGSFSPKS 1199
1201 SSGCSSKTEHMSSSSENTSTPSSNAKRONQLHPSSAQSRSGQPKAATCKVSSINQETIQ 1260
1200 SSGCSSKTEHMSSSSENTSTPSSNAKRONQLHPSSAQSRSGQPKAATCKVSSINQETIQ 1259
1261 TYCVEDTPICFRSCSSLSLSSAEDIGCNQTTQEADSAANTLQIAIEKIGTRSAEDPV 1320
1260 TYCVEDTPICFRSCSSLSLSSAEDIGCNQTTQEADSAANTLQIAIEKIGTRSAEDPV 1319

```

1321 SEVPVAVQHPRKTSRLOQSSLSSEGAHKAFFSSGAKSPSKGAQTPKSPPEHYVQET 1380  
1320 SEVPVAVQHPRKTSRLOQSSLSSEGAHKAFFSSGAKSPSKGAQTPKSPPEHYVQET 1379  
1381 PLMFSTRCTSVSSLDSPESRISASVQPCSGMVSGIISPSDLDPSPGOTMPPSRSKTPP 1440  
1380 PLMFSTRCTSVSSLDSPESRISASVQPCSGMVSGIISPSDLDPSPGOTMPPSRSKTPP 1439  
1441 PPQTAQTKREVKNKAPTAKESEGPQAAVNAVQVQVLPDADTLHFAATESPDGF 1500  
1440 PPQTAQTKREVKNKAPTAKESEGPQAAVNAVQVQVLPDADTLHFAATESPDGF 1499  
1501 SCSSLSALSALDPFFQKQVLELIMPPVQENDNGNETSEFQPKESNENQEKAEKTI DSE 1560  
1500 SCSSLSALSALDPFFQKQVLELIMPPVQENDNGNETSEFQPKESNENQEKAEKTI DSE 1559  
1561 KDLDDSDDDDDIIEILBECIISAMFTKSSRKAKPAQATASKLPPPVARKPSQLPVYKLLPS 1620  
1560 KDLDDSDDDDDIIEILBECIISAMFTKSSRKAKPAQATASKLPPPVARKPSQLPVYKLLPS 1619  
1621 QNLQPKQVSVFTPGDDMPRVYCEGTPINFSTATSLDITTESPNELAAEGVRGGAQ 1680  
1620 QNLQPKQVSVFTPGDDMPRVYCEGTPINFSTATSLDITTESPNELAAEGVRGGAQ 1679  
1681 SGFEKRDITPTBGRSTDEAQGGKTSVITPELDDNKABEGDILACINSAMPKGSHPK 1740  
1680 SGFEKRDITPTBGRSTDEAQGGKTSVITPELDDNKABEGDILACINSAMPKGSHPK 1739  
1741 FRVKIMQVQOASASSAPKNQOLGKKKTPSPVKPIQNTTEYTRVRKNADSKNLN 1800  
1740 FRVKIMQVQOASASSAPKNQOLGKKKTPSPVKPIQNTTEYTRVRKNADSKNLN 1799  
1801 AERFVSNKDKSKONLKNNSKDFNDKLNNEDRVKGSFAFDPSPHHVTPTEGTPYCFERN 1860  
1800 AERFVSNKDKSKONLKNNSKDFNDKLNNEDRVKGSFAFDPSPHHVTPTEGTPYCFERN 1859  
1861 SLSSLDFFDDDDVLSREKAEKRAKAKENKSEAKVTSHTELTSNQOSANKTQAIKQPINR 1920  
1860 SLSSLDFFDDDDVLSREKAEKRAKAKENKSEAKVTSHTELTSNQOSANKTQAIKQPINR 1919  
1921 GQPKPILQKOSTPQSSKOIPDRGAATDEKLNFAIENTPVCFESHNSLSLSDIDQENN 1980  
1920 GQPKPILQKOSTPQSSKOIPDRGAATDEKLNFAIENTPVCFESHNSLSLSDIDQENN 1979  
1981 NKNEPIKETEPDPSQEPSPKQASGYAPKSFHVEDTPVCFERNSSLSLSDIDQENLLQ 2040  
1980 NKNEPIKETEPDPSQEPSPKQASGYAPKSFHVEDTPVCFERNSSLSLSDIDQENLLQ 2039  
2041 ECISAMPKPKKPSRLKGDNEKHSRNMGGIIGEDTLDLKOIQRDPDSEHGLSPDSENF 2100  
2040 ECISAMPKPKKPSRLKGDNEKHSRNMGGIIGEDTLDLKOIQRDPDSEHGLSPDSENF 2099  
2101 WKAIQEGANSIVSLHQAALCLSCQASDSDSILSLKSGISLQSPFFHLTPDQEKPT 2160  
2100 WKAIQEGANSIVSLHQAALCLSCQASDSDSILSLKSGISLQSPFFHLTPDQEKPT 2159  
2161 SNKGPRILKPGKSTLETTKKIESEKGIKGGKVKYKSLITGKVRNSSEISGQWKQPLQAN 2220  
2160 SNKGPRILKPGKSTLETTKKIESEKGIKGGKVKYKSLITGKVRNSSEISGQWKQPLQAN 2219  
2221 MPISIRGRTHIHPGVNRSSTSPVSKGPPKLTPTASKPSBEGOTATTSPRGAKPSVKS 2280  
2220 MPISIRGRTHIHPGVNRSSTSPVSKGPPKLTPTASKPSBEGOTATTSPRGAKPSVKS 2279  
2281 ELSFVARQTSIQGSSKAPSRGSRDSTPSRPAQQLSRPIQSPGRNISPGRNGISPPN 2340  
2280 ELSFVARQTSIQGSSKAPSRGSRDSTPSRPAQQLSRPIQSPGRNISPGRNGISPPN 2339  
2341 KLSQLPRTSPSTASTKSSGSMYSYTSRQWQONLTKOTGLSKNASSIPRSESASKG 2400  
2340 KLSQLPRTSPSTASTKSSGSMYSYTSRQWQONLTKOTGLSKNASSIPRSESASKG 2399

2401 LQNMNNGANKVELSRMSTKSSGSESDSRPVLVRSQTFIKEAPSPTLRKLBEA 2460  
2400 LQNMNNGANKVELSRMSTKSSGSESDSRPVLVRSQTFIKEAPSPTLRKLBEA 2459  
2461 SFESLSPSRSPASPTRSQATPVLSPSLPDMSLSTHSSVQAGWRKLPPLSPPTIENDG 2520  
2460 SFESLSPSRSPASPTRSQATPVLSPSLPDMSLSTHSSVQAGWRKLPPLSPPTIENDG 2519  
2521 RPAKHDIARSHSPSRPLPINRSGTWKREHKSUSLSIPRVSTWRTGSSSSILSASSES 2580  
2520 RPAKHDIARSHSPSRPLPINRSGTWKREHKSUSLSIPRVSTWRTGSSSSILSASSES 2579  
2581 SEKAKSEDEKHVNSISGTKQSKENQVSAKGTWRIKENEPSTNSTQTSVSSGATNGAES 2640  
2580 SEKAKSEDEKHVNSISGTKQSKENQVSAKGTWRIKENEPSTNSTQTSVSSGATNGAES 2639  
2641 KTLIYQMAPAVSKTEDVWVRIEDCPINNPRSGRPTGNTPPVIDSVSEKANPNIKDSKN 2700  
2640 KTLIYQMAPAVSKTEDVWVRIEDCPINNPRSGRPTGNTPPVIDSVSEKANPNIKDSKN 2699  
2701 QAKQNVGNGSVPMETVGLNELNSFIOVDAPDQKGTIKQCONNPVSVSETNESSIVERT 2760  
2700 QAKQNVGNGSVPMETVGLNELNSFIOVDAPDQKGTIKQCONNPVSVSETNESSIVERT 2759  
2761 PFSSSSSKHSPSGTVAARVTPFNYPNPSPKSSADSTSPSQIPTPVNNNTKKRDSKT 2820  
2760 PFSSSSSKHSPSGTVAARVTPFNYPNPSPKSSADSTSPSQIPTPVNNNTKKRDSKT 2819  
2821 DSTESSGTQSPKHSGLVTSV 2843  
2820 DSTESSGTQSPKHSGLVTSV 2842

## RESULT 11

US-08-449-731-7

; Sequence 7, Application US/08449731

; Patent No. 6413727

; GENERAL INFORMATION:

; APPLICANT: ALBERTSEN, HANS

; ANAND, RAKESH

; CARLSON, MARY

; GRODEN, JOANNA

; HEDGE, PHILIP J.

; JOSLYN, GEOFF

; KINZLER, KENNETH

; MARKHAM, ALEXANDER F.

; NAKAMURA, YUSUKE

; THLIVERIS, ANDREW

; TITLE OF INVENTION: INHERITED AND SOMATIC MUTATIONS OF APC

; GENE IN COLORECTAL CANCER IN HUMANS

; NUMBER OF SEQUENCES: 102

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Banner &amp; Allegretti, LTD

; STREET: 1001 G Street, NW

; CITY: Washington

; STATE: D.C.

; COUNTRY: USA

; ZIP: 20001-4598

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentin Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/449,731

; FILING DATE: 25-May-1995

; CLASSIFICATION: &lt;Unknown&gt;

; PRIORITY INFORMATION:

; APPLICATION NUMBER: 08/289,548

; FILING DATE: 12-AUG-1994

; ATTORNEY/AGENT INFORMATION:

; NAME: Kagan, Sarah A.

; REGISTRATION NUMBER: 32,141



REFERENCE/DOCKET NUMBER: 1107.46943  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-508-9100  
TELEFAX: 202-508-9299  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2842 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
ORIGINAL SOURCE:  
ORGANISM: Homo sapiens  
IMMEDIATE SOURCE:  
CLONE: APC  
SEQUENCE DESCRIPTION: SEQ ID NO: 7:  
US-08-449-731-7

Query Match 99.9%; Score 14548.5; DB 4; Length 2842;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2842; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

Qy	1	MAAASVDOLKKOVALKVENSLRQLEEDNSNHLTKLEASNMKEVLKQGGSTEDAM	60
Db	1	MAAASVDOLKKOVALKVENSLRQLEEDNSNHLTKLEASNMKEVLKQGGSTEDAM	60
Qy	61	ASSGGIDLLERLKEINLDPSSNPPGVKLRSKMSLSRSGREGSVSRSGECSPVPMGSPFR	120
Db	61	ASSGGIDLLERLKEINLDPSSNPPGVKLRSKMSLSRSGREGSVSRSGECSPVPMGSPFR	120
Qy	121	RGVNGSRSTGYLSELEKERSLLADLDKBEKXDWTYAOIUNTKRIDSPLTENPSL	180
Db	121	RGVNGSRSTGYLSELEKERSLLADLDKBEKXDWTYAOIUNTKRIDSPLTENPSL	179
Qy	181	QDMDTRRQLEYEARQIRVAMEBOLGTCODMEKRAQRRIARQQEKDILRLRQLQSOAT	240
Db	180	QDMDTRRQLEYEARQIRVAMEBOLGTCODMEKRAQRRIARQQEKDILRLRQLQSOAT	239
Qy	241	EAERSSQNHGTGSHDAERNGQGVGEINNATSGNGQGGTTRMDHETASVLSSTHSA	300
Db	240	EAERSSQNHGTGSHDAERNGQGVGEINNATSGNGQGGTTRMDHETASVLSSTHSA	299
Qy	301	PRRLTSHLGTKVEMYVLSLMLGTHDKDMSRTLLAMSSODSCISMEQSGCLPILIQLL	360
Db	300	PRRLTSHLGTKVEMYVLSLMLGTHDKDMSRTLLAMSSODSCISMEQSGCLPILIQLL	359
Qy	361	HGNDKXSVLLGNSRGSKEARAFASAAALHNIHSQPDQDKRGRREIRVLHLLQIRAYCETC	420
Db	360	HGNDKXSVLLGNSRGSKEARAFASAAALHNIHSQPDQDKRGRREIRVLHLLQIRAYCETC	419
Qy	421	NEWQBAHEFGMDQDQXNPAPVVEHQICPAVCVLMKLSFDEEHRHAMNELGGLOIAELLQ	480
Db	420	NEWQBAHEFGMDQDQXNPAPVVEHQICPAVCVLMKLSFDEEHRHAMNELGGLOIAELLQ	479
Qy	481	VDCEMYGLTNDHYSITLRRYAGMALTNLTFGDVANKATLCMKGCMRALVAQLKSESDL	540
Db	480	VDCEMYGLTNDHYSITLRRYAGMALTNLTFGDVANKATLCMKGCMRALVAQLKSESDL	539
Qy	541	QOVIASVLRLNSWRADVNSKTLRLRVGSVKALMECALEVYKKESTLKSVLNALWNLSAHCT	600
Db	540	QOVIASVLRLNSWRADVNSKTLRLRVGSVKALMECALEVYKKESTLKSVLNALWNLSAHCT	599
Qy	601	ENKADICAVDQALAFVLGTLTVRSOTNTLAIIESGGGILRVNSSLIATNEDHROLRENN	660
Db	600	ENKADICAVDQALAFVLGTLTVRSOTNTLAIIESGGGILRVNSSLIATNEDHROLRENN	659
Qy	661	CIQTLLOHLKSHSLTVSNACGTLNLSARNPKDQALWDMGAVSMKLNLIHSHKMIAM	720
Db	660	CIQTLLOHLKSHSLTVSNACGTLNLSARNPKDQALWDMGAVSMKLNLIHSHKMIAM	719
Qy	721	GSAALRLNLMANRPAPYKXDNIMSPGSSILPSLHVKKQKALELDAQHLSLTFNIDNLS	780
Db	720	GSAALRLNLMANRPAPYKXDNIMSPGSSILPSLHVKKQKALELDAQHLSLTFNIDNLS	779

Qy	781	PKASHRSKQRHKOSLYGDYVFDNRRHDDNRSDNFTGNMTVLSPYLNTTVLPSSSSSRGS	840
Db	780	PKASHRSKQRHKOSLYGDYVFDNRRHDDNRSDNFTGNMTVLSPYLNTTVLPSSSSSRGS	839
Qy	841	LDSSRSSEKDRSLERERIGLGNYPATENPGTSSKEGLQISTTAAQIAKVMEEVSAIHTS	900
Db	840	LDSSRSSEKDRSLERERIGLGNYPATENPGTSSKEGLQISTTAAQIAKVMEEVSAIHTS	899
Qy	901	QEDRSSTTELHCVTDERNALRRSSAAATHSNYNTFTKSENENRNTCSMPYAKLEYKRSS	960
Db	900	QEDRSSTTELHCVTDERNALRRSSAAATHSNYNTFTKSENENRNTCSMPYAKLEYKRSS	959
Qy	961	NDLSNSVSSSDGYGKRGOMKPSIESYSEDDSEKFCYGOVPADLAHLKIHSANEMDNDGE	1020
Db	960	NDLSNSVSSSDGYGKRGOMKPSIESYSEDDSEKFCYGOVPADLAHLKIHSANEMDNDGE	1019
Qy	1021	LDPINYSKYSDQNLNSGQSPSONERWARPKHIIIEDEIKQSEQRNQSTTYPVYTE	1080
Db	1020	LDPINYSKYSDQNLNSGQSPSONERWARPKHIIIEDEIKQSEQRNQSTTYPVYTE	1079
Qy	1081	STDDKHLKQPHFGQCECVSPYSRGANGSETNRVGSNHGINONVQSLOEDDYEDDKP	1140
Db	1080	STDDKHLKQPHFGQCECVSPYSRGANGSETNRVGSNHGINONVQSLOEDDYEDDKP	1139
Qy	1141	TNYSERYSEEEQHEEERPTNYSIKYNEEKRHVDQIDYSLKYATDIPSSQKQSFSPKS	1200
Db	1140	TNYSERYSEEEQHEEERPTNYSIKYNEEKRHVDQIDYSLKYATDIPSSQKQSFSPKS	1199
Qy	1201	SSQSSKTEHMSSESTSTPSSNAKRONQLHPSSAQSRSGQPKAATCKVSSINQETIQ	1260
Db	1200	SSQSSKTEHMSSESTSTPSSNAKRONQLHPSSAQSRSGQPKAATCKVSSINQETIQ	1259
Qy	1261	TYCVEDTPICFSCSSLSLSAEDIGCNOTTQEADSNANTLQIAIEIKIGTRSAEDPV	1320
Db	1260	TYCVEDTPICFSCSSLSLSAEDIGCNOTTQEADSNANTLQIAIEIKIGTRSAEDPV	1319
Qy	1321	SEVPASVQHPRTKSRILQSSLSSESARHKAVFSSCAKSPSKGACTPKGPPHYVQET	1380
Db	1320	SEVPASVQHPRTKSRILQSSLSSESARHKAVFSSCAKSPSKGACTPKGPPHYVQET	1379
Qy	1381	PLMFSRCTSVSSLDSPESRSIASVQSEPCSGMVSGIISPSDLPDPSGQTPMPSRKTTP	1440
Db	1380	PLMFSRCTSVSSLDSPESRSIASVQSEPCSGMVSGIISPSDLPDPSGQTPMPSRKTTP	1439
Qy	1441	PPQTAQTKREVPKNAKPTAKRESQPKQAAVNAVQVQVLPDADTLHLHFAESTPDGF	1500
Db	1440	PPQTAQTKREVPKNAKPTAKRESQPKQAAVNAVQVQVLPDADTLHLHFAESTPDGF	1499
Qy	1501	SCSSLSALSIDBFFIQKQVLELIMPPVQENDNGNETESEQPKESNENQEKAEKTIDSE	1560
Db	1500	SCSSLSALSIDBFFIQKQVLELIMPPVQENDNGNETESEQPKESNENQEKAEKTIDSE	1559
Qy	1561	KDLDSDSDDDIIEILBECIISAMPTKSSRKAKKPAOTASKLPPPVARKPSQLPVYKLLPS	1620
Db	1560	KDLDSDSDDDIIEILBECIISAMPTKSSRKAKKPAOTASKLPPPVARKPSQLPVYKLLPS	1619
Qy	1621	QNRLOPKHVSFTPGDDMPRVYCVGEGTPIINFSTATSLSDLTIESPPNELAAGVGVRGAQ	1680
Db	1620	QNRLOPKHVSFTPGDDMPRVYCVGEGTPIINFSTATSLSDLTIESPPNELAAGVGVRGAQ	1679
Qy	1681	SGFEKEDTITPTGRSTDEAQGGKTSVTLPELDDNKAEBEGDILAEICINSAMPKGSHKP	1740
Db	1680	SGFEKEDTITPTGRSTDEAQGGKTSVTLPELDDNKAEBEGDILAEICINSAMPKGSHKP	1739
Qy	1741	FRVKIMQVQOASASSAPNKNQDQKKKPTSPVKPIPQNTTEYTRVRKNADSKNNLN	1800
Db	1740	FRVKIMQVQOASASSAPNKNQDQKKKPTSPVKPIPQNTTEYTRVRKNADSKNNLN	1799
Qy	1801	AERFSDNKKONLKNNSKDFNDKLPNNEDVRGSAFDSPHYTPTEGTTCYCSRND	1860
Db	1800	AERFSDNKKONLKNNSKDFNDKLPNNEDVRGSAFDSPHYTPTEGTTCYCSRND	1859

QY 1861 SLGLDDEDDVLSRKAELRKAENKESAKVTSHTLTSTNOQANKTOAIKQPINR 1920  
 Db 1860 SLGLDDEDDVLSRKAELRKAENKESAKVTSHTLTSTNOQANKTOAIKQPINR 1919  
 QY 1921 GQPKILQKSTPOSSKDTPDGAATDEKLOFPAIENTVCFSHNSLSLSDIDQENN 1980  
 Db 1920 GQPKILQKSTPOSSKDTPDGAATDEKLOFPAIENTVCFSHNSLSLSDIDQENN 1979  
 QY 1981 NKNEPIKETEPDQSGEPKQASGYPKSFVVEDTPVCFSRNSLSLSDIDQENN 2040  
 Db 1980 NKNEPIKETEPDQSGEPKQASGYPKSFVVEDTPVCFSRNSLSLSDIDQENN 2039  
 QY 2041 ECISAPKPKKPSRLKGNKESPRNMGILGEDTLDKDTQRPDSEHGLSPDSENF 2100  
 Db 2040 ECISAPKPKKPSRLKGNKESPRNMGILGEDTLDKDTQRPDSEHGLSPDSENF 2099  
 QY 2101 WKATQEGANSIVSLHQAACLSRQASDSDSILSLKSGISLGSPFHLTPDQEKPT 2160  
 Db 2100 WKATQEGANSIVSLHQAACLSRQASDSDSILSLKSGISLGSPFHLTPDQEKPT 2159  
 QY 2161 SNKGPRILKPEKSTLTETKIEBSKIGKGGKVKYKSLITGKVRNSSEISGQWKQPLQAN 2220  
 Db 2160 SNKGPRILKPEKSTLTETKIEBSKIGKGGKVKYKSLITGKVRNSSEISGQWKQPLQAN 2219  
 QY 2221 MPSISRGRTMIHIGVNSSSSTSPVSKGPPPLKTPASKPSGQATTSRCAKPSVKS 2280  
 Db 2220 MPSISRGRTMIHIGVNSSSSTSPVSKGPPPLKTPASKPSGQATTSRCAKPSVKS 2279  
 QY 2281 ELSVARTQSIGSSKAPSRGSRDTPGRPAQPLSRPIQSPGRNSISPGRNGISPEN 2340  
 Db 2280 ELSVARTQSIGSSKAPSRGSRDTPGRPAQPLSRPIQSPGRNSISPGRNGISPEN 2339  
 QY 2341 KLSQLPSTSPSTASTSSGSGKMSYTSRGRQSQNLTKQTGLSKNASSIPRSESASG 2400  
 Db 2340 KLSQLPSTSPSTASTSSGSGKMSYTSRGRQSQNLTKQTGLSKNASSIPRSESASG 2399  
 QY 2401 LQMNNGNGANKKVELSRMSTKSSGSDRSPVLRVOSTTKEAPSPTLRKLKLESA 2460  
 Db 2400 LQMNNGNGANKKVELSRMSTKSSGSDRSPVLRVOSTTKEAPSPTLRKLKLESA 2459  
 QY 2461 SFESLSPSPSPSTRQAOQTPVLSPLDMSLSTHSSVQAGWRKLPNLSPTIENDG 2520  
 Db 2460 SFESLSPSPSPSTRQAOQTPVLSPLDMSLSTHSSVQAGWRKLPNLSPTIENDG 2519  
 QY 2521 RPAKRHDIASSHSPRLPINSRGTWKRHSKSSSLPRVSTWRTGSSSSTLSASSES 2580  
 Db 2520 RPAKRHDIASSHSPRLPINSRGTWKRHSKSSSLPRVSTWRTGSSSSTLSASSES 2579  
 QY 2581 SEKAKSEDEKXVNSISGTQSKENQVSAKTWKIKENEPSPNTSTQTVSSGATNGAES 2640  
 Db 2580 SEKAKSEDEKXVNSISGTQSKENQVSAKTWKIKENEPSPNTSTQTVSSGATNGAES 2639  
 QY 2641 KTLIYQAPAVSKTDEYVWRIEDCPINPRSGSPGTGTPPVTDVSEKANPNIKDSKN 2700  
 Db 2640 KTLIYQAPAVSKTDEYVWRIEDCPINPRSGSPGTGTPPVTDVSEKANPNIKDSKN 2699  
 QY 2701 QAKQNVGSGVPMRTVGLNRLNSFIQVADPQKGTIEKPGQNNPVPVSETNESSIVERT 2760  
 Db 2700 QAKQNVGSGVPMRTVGLNRLNSFIQVADPQKGTIEKPGQNNPVPVSETNESSIVERT 2759  
 QY 2761 PFSSSSSKHSSPGCTVAARVTPNPNPSRKSADSTSRPSQIPTPVNNNTKEDSKT 2820  
 Db 2760 PFSSSSSKHSSPGCTVAARVTPNPNPSRKSADSTSRPSQIPTPVNNNTKEDSKT 2819  
 QY 2821 DSTESSGTQSPKRGHSGSYLTVS 2843  
 Db 2820 DSTESSGTQSPKRGHSGSYLTVS 2842

GENERAL INFORMATION:  
 APPLICANT: ALBERTSEN, HANS  
 APPLICANT: ANAND, RAKESH  
 APPLICANT: CARLSON, MARY  
 APPLICANT: GRODEN, JOANNA  
 APPLICANT: HEDGE, PHILIP J.  
 APPLICANT: JOSLYN, GEOFF  
 APPLICANT: KINZLER, KENNETH  
 APPLICANT: MARKHAM, ALEXANDER F.  
 APPLICANT: NAKAMURA, YUSUKE  
 APPLICANT: TELIVERIS, ANDREW  
 TITLE OF INVENTION: INHERITED AND SOMATIC MUTATIONS OF APC  
 TITLE OF INVENTION: GENE IN COLORECTAL CANCER IN HUMANS  
 NUMBER OF SEQUENCES: 94  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Banner, Birch, McKie & Beckett  
 STREET: 1001 G Street, NW  
 CITY: Washington  
 STATE: D.C.  
 COUNTRY: USA  
 ZIP: 20001-4598  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA: US/07/741,940  
 APPLICATION NUMBER: 32,141  
 FILING DATE: 19920109  
 CLASSIFICATION: 435  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Kagan, Sarah A.  
 REGISTRATION NUMBER: 1107.035574  
 REFERENCE/DOCKET NUMBER: 1107.035574  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 202-508-9100  
 TELEFAX: 202-508-9299  
 INFORMATION FOR SEQ ID NO: 2:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 2843 amino acids  
 TYPE: AMINO ACID  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 US-07-741-940-2

Query Match 99.8%; Score 14533; DB 1; Length 2843;  
 Best Local Similarity 99.8%; Pred. No. 0;  
 Matches 2836; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 MAASVDQLLKQVAKALXNENLNRLQEDNSNHLTKLETSANMKVILKQLQGSIEDEAM 60  
 Db 1 MAASVDQLLKQVAKALXNENLNRLQEDNSNHLTKLETSANMKVILKQLQGSIEDEAM 60  
 QY 61 ASSGQIDLLLELKEINLDSSNFPKVLRSKMSLSYSGRSGSVSSRSGECSVPVPMGSPFR 120  
 Db 61 ASSGQIDLLLELKEINLDSSNFPKVLRSKMSLSYSGRSGSVSSRSGECSVPVPMGSPFR 120  
 QY 121 RGFVNGSRESTGYLEELKERSLLADLDKEEKDWMYLAQLQNLTKRIDSLPTENFSL 180  
 Db 121 RGFVNGSRESTGYLEELKERSLLADLDKEEKDWMYLAQLQNLTKRIDSLPTENFSL 180  
 QY 181 QTDWTRQLEAYEQIRVAMEEQIGTQDMKRAQRRIARIQQIEKDILRIROLQSQAT 240  
 Db 181 QTDWTRQLEAYEQIRVAMEEQIGTQDMKRAQRRIARIQQIEKDILRIROLQSQAT 240  
 QY 241 EAERSSQNKHETGSHDAERQNEGQVGEINMATSGNGQSTTRMDHETASVLSSTHSA 300  
 Db 241 EAERSSQNKHETGSHDAERQNEGQVGEINMATSGNGQSTTRMDHETASVLSSTHSA 300  
 QY 301 PRLTSHLGTVMYVYLLSMLGTHDKDMSRTLLAMSSSDSCISMRQSGCLPLLIQLL 360  
 Db 301 PRLTSHLGTVMYVYLLSMLGTHDKDMSRTLLAMSSSDSCISMRQSGCLPLLIQLL 360

QY 361 HGNDKDSVLLGNSRGSKARARASAAALHNIHSQDDPKRGREIRVLHLBOIRAYCETC 420  
DB 361 HGNDKDSVLLGNSRGSKARARASAAALHNIHSQDDPKRGREIRVLHLBOIRAYCETC 420  
QY 421 WUQAEHFGMDQDNMPAPVEHQICPAVCVLMKLSFDEEHRHAMNELGGIQAIAELLQ 480  
DB 421 WUQAEHFGMDQDNMPAPVEHQICPAVCVLMKLSFDEEHRHAMNELGGIQAIAELLQ 480  
QY 481 VDCWYGLTNDHYSITLRYAGWALTNLTFPGVANKATILCSMKGCWRAIVAOIKSESD 540  
DB 481 VDCWYGLTNDHYSITLRYAGWALTNLTFPGVANKATILCSMKGCWRAIVAOIKSESD 540  
QY 541 QOVIASVLNLSWRADVNSKTLREVGSKVLMKALEVKKSESTLKSIVLSALWNLISACT 600  
DB 541 QOVIASVLNLSWRADVNSKTLREVGSKVLMKALEVKKSESTLKSIVLSALWNLISACT 600  
QY 601 ENKADICAVDGAFLAVGLTTRYSQNTLAIIESGGILRNVSIIATNEDHRQILRENN 660  
DB 601 ENKADICAVDGAFLAVGLTTRYSQNTLAIIESGGILRNVSIIATNEDHRQILRENN 660  
QY 661 CIQTLLOHLKSHSLATVSNACGTLNLSARNPKDOEALWDMGAVSNLXNLHSHKMIAM 720  
DB 661 CIQTLLOHLKSHSLATVSNACGTLNLSARNPKDOEALWDMGAVSNLXNLHSHKMIAM 720  
QY 721 GSAALRNLMANRPKYKDANIMSPGSSLPSSHVKQKALEAELDAOHLSETFDNIDLS 780  
DB 721 GSAALRNLMANRPKYKDANIMSPGSSLPSSHVKQKALEAELDAOHLSETFDNIDLS 780  
QY 781 PKASHRSKORHKQSLYGVDFDTNRHDDNRSNFTGNMTVLSPLYLNTTLPSSSSSRGS 840  
DB 781 PKASHRSKORHKQSLYGVDFDTNRHDDNRSNFTGNMTVLSPLYLNTTLPSSSSSRGS 840  
QY 841 LPSRSSEKDRSLERERGI GLGNVHPATENPGTSSKRGQIQTAAQIAKVMEEVSAIHTS 900  
DB 841 LPSRSSEKDRSLERERGI GLGNVHPATENPGTSSKRGQIQTAAQIAKVMEEVSAIHTS 900  
QY 901 QEDRSSGSTTEHLCHVTDERNALRRSAATHNTYNTFKSENSNRTCSMPYAKLEYKSS 960  
DB 901 QEDRSSGSTTEHLCHVTDERNALRRSAATHNTYNTFKSENSNRTCSMPYAKLEYKSS 960  
QY 961 NDLSNVSSDGYGKRGQMKPIESYSEDESKFCYGOYPADLAHKIHSANHMDNDGE 1020  
DB 961 NDLSNVSSDGYGKRGQMKPIESYSEDESKFCYGOYPADLAHKIHSANHMDNDGE 1020  
QY 1021 LPTPINYSIKYSDQLNSGRQSPQNERWARPKHIEIEIKOSEQRQRNOSTTYPVYTE 1080  
DB 1021 LPTPINYSIKYSDQLNSGRQSPQNERWARPKHIEIEIKOSEQRQRNOSTTYPVYTE 1080  
QY 1081 STDDKHLKFPQHFQGOECVSPYRSRGANGSETNRYGNSHGINQVNSQSI-COEDDYDDKP 1140  
DB 1081 STDDKHLKFPQHFQGOECVSPYRSRGANGSETNRYGNSHGINQVNSQSI-COEDDYDDKP 1140  
QY 1141 TNYSEYSEEEQHEBERPTNYSIKYNEBKRVDPIDYSLKYATDIPSSQKQSFSSKS 1200  
DB 1141 TNYSEYSEEEQHEBERPTNYSIKYNEBKRVDPIDYSLKYATDIPSSQKQSFSSKS 1200  
QY 1201 SSGQSKTEHMSSENSTTPSNAKRONQLHPSSNAQSRSGOPOKAATCKVSSINQETIQ 1260  
DB 1201 SSGQSKTEHMSSENSTTPSNAKRONQLHPSSNAQSRSGOPOKAATCKVSSINQETIQ 1260  
QY 1261 TVCVEDTPICFRCSLSLSAEDBIGCNOTTOBADSANTLQIAEIEIKIGITRAEDPV 1320  
DB 1261 TVCVEDTPICFRCSLSLSAEDBIGCNOTTOBADSANTLQIAEIEIKIGITRAEDPV 1320  
QY 1321 SEVPVAVSQHPRTKSSRLOGSSLSSESARHAKVEFPSSGAKSPKSGAQTPKSPPEHYVQET 1380  
DB 1321 SEVPVAVSQHPRTKSSRLOGSSLSSESARHAKVEFPSSGAKSPKSGAQTPKSPPEHYVQET 1380  
QY 1381 PLMFBSCTSVSSLDSPESRISASSVQSEPCSGWVGIIISPSDLPSPGQTMPPSSKTPP 1440  
DB 1381 PLMFBSCTSVSSLDSPESRISASSVQSEPCSGWVGIIISPSDLPSPGQTMPPSSKTPP 1440  
QY 1441 PPPQTAQTKREVFNKAPTAERESGPKQAANAAVQVQLPDADTLLHFATESTPDGF 1500

DB 1441 PPPQTAQTKREVFNKAPTAERESGPKQAANAAVQVQLPDADTLLHFATESTPDGF 1500  
QY 1501 SCSSLSALSILDEPFFIQKQVELRIMPVQENDNGNETESQPKESNENQKEAKTIDSE 1560  
DB 1501 SCSSLSALSILDEPFFIQKQVELRIMPVQENDNGNETESQPKESNENQKEAKTIDSE 1560  
QY 1561 KDLDDSDDDDDIETLBECHISAMPTSSRKAKKPAQATASKLPVPVARKPQOLPVYKLLPS 1620  
DB 1561 KDLDDSDDDDDIETLBECHISAMPTSSRKAKKPAQATASKLPVPVARKPQOLPVYKLLPS 1620  
QY 1621 QNRLOPKHVSFTPGDDMPRVYCVEGTPIINFSTATSLSDLTIESPPNELAAGSVRGGAQ 1680  
DB 1621 QNRLOPKHVSFTPGDDMPRVYCVEGTPIINFSTATSLSDLTIESPPNELAAGSVRGGAQ 1680  
QY 1681 SGEFEKEDTTPTEGRSTDEAOGGKTSSVTPELDDNKABEGDILAEICINAMPKGSHKP 1740  
DB 1681 SGEFEKEDTTPTEGRSTDEAOGGKTSSVTPELDDNKABEGDILAEICINAMPKGSHKP 1740  
QY 1741 FRVKKINDQVQOASASSAPNKNQLOGKKKPTSPVKPIPONTTEYRTRVRKNADSKNNLN 1800  
DB 1741 FRVKKINDQVQOASASSAPNKNQLOGKKKPTSPVKPIPONTTEYRTRVRKNADSKNNLN 1800  
QY 1801 AERVFSNKNKONLKNKSKDFNDKLPNNEDRVRSFAFDSPHHYTPIEGTYPYCFSRND 1860  
DB 1801 AERVFSNKNKONLKNKSKDFNDKLPNNEDRVRSFAFDSPHHYTPIEGTYPYCFSRND 1860  
QY 1861 SLSSLDPDDDDVLSREKAEALRKAENKSEAKVTSHTELTSNQOSANKTQAIKQPINR 1920  
DB 1861 SLSSLDPDDDDVLSREKAEALRKAENKSEAKVTSHTELTSNQOSANKTQAIKQPINR 1920  
QY 1921 GQPKPILQKOSTPPQSSKDIIPDRGAATDEKLONPATIENTPVCFSHNSLSLSDIDQENN 1980  
DB 1921 GQPKPILQKOSTPPQSSKDIIPDRGAATDEKLONPATIENTPVCFSHNSLSLSDIDQENN 1980  
QY 1981 NKNEPIKETEPDPSQGEPSKQASGYAPKSFVEVETPVCFSNSSLSSLSISIDSDLLQ 2040  
DB 1981 NKNEPIKETEPDPSQGEPSKQASGYAPKSFVEVETPVCFSNSSLSSLSISIDSDLLQ 2040  
QY 2041 ECISAMPKPKKPSRLKGDNEKHSRNMGGILGEBDLTDLKDQRPDSEHGLSPDSNFD 2100  
DB 2041 ECISAMPKPKKPSRLKGDNEKHSRNMGGILGEBDLTDLKDQRPDSEHGLSPDSNFD 2100  
QY 2101 WKAIQEGANSIVSSLHQAACCLSRQASSDSDSILSKGSIILGSPFHITPDQEEKPT 2160  
DB 2101 WKAIQEGANSIVSSLHQAACCLSRQASSDSDSILSKGSIILGSPFHITPDQEEKPT 2160  
QY 2161 SNKGPRILKPGKSTLTETKIESESGIKGKKVYKSLITGKVRNSSEISGQMKQPLQAN 2220  
DB 2161 SNKGPRILKPGKSTLTETKIESESGIKGKKVYKSLITGKVRNSSEISGQMKQPLQAN 2220  
QY 2221 MPSTSRGRTMIHIPGVNNSSTSPVSKGPPKLTTPASKSPSEGQTATTSPRGAKPSVK 2280  
DB 2221 MPSTSRGRTMIHIPGVNNSSTSPVSKGPPKLTTPASKSPSEGQTATTSPRGAKPSVK 2280  
QY 2281 ELSVAREQTSQICGSSKAPSRGSRDSTPSRPAQOPLSRPIQSPGRNSISPGRNGISPPN 2340  
DB 2281 ELSVAREQTSQICGSSKAPSRGSRDSTPSRPAQOPLSRPIQSPGRNSISPGRNGISPPN 2340  
QY 2341 KLSQLPRTSPSTASTKSSGSGKMSYTPSGRQMSQQNLTKQTGLSKNASSIPRESASKG 2400  
DB 2341 KLSQLPRTSPSTASTKSSGSGKMSYTPSGRQMSQQNLTKQTGLSKNASSIPRESASKG 2400  
QY 2401 LNQWNGNGANKVELSRMSSTKSSGSGKMSYTPSGRQMSQQNLTKQTGLSKNASSIPRESASKG 2460  
DB 2401 LNQWNGNGANKVELSRMSSTKSSGSGKMSYTPSGRQMSQQNLTKQTGLSKNASSIPRESASKG 2460  
QY 2461 SFESLSPSSRPASPTRQAOTPVLSPLPDMSLSTHSSVQAGWRKLPPLNLSPTIEYNDG 2520  
DB 2461 SFESLSPSSRPASPTRQAOTPVLSPLPDMSLSTHSSVQAGWRKLPPLNLSPTIEYNDG 2520  
QY 2521 RPAKXHDIAKSHSPSRPLPINRSGTWKRBHSHKSSSLPRVSTWRTGRTSSSSILSASSES 2580

Db 2521 RPAKRHDIAKSHSPRLPINRSGTWKREHSHSSSLPRVSTWRRTGSSSSILSASSPS 2580  
Qy 2581 SEKAKSEDEKXVNSISGTQSKENQVSAKGTWRKIKENEFSPNTSQTSSGATNGAES 2640  
Db 2581 SEKAKSEDEKXVNSISGTQSKENQVSAKGTWRKIKENEFSPNTSQTSSGATNGAES 2640  
Qy 2641 KTLIYQAPAVSKTEDVWVRIEDCPINNPRSGRSPGTNTPPVIDSVSEKANPNIKDSKN 2700  
Db 2641 KTLIYQAPAVSKTEDVWVRIEDCPINNPRSGRSPGTNTPPVIDSVSEKANPNIKDSKN 2700  
Qy 2701 QAKONVNGSVPMRTVGLNRLKNSFIQVADPDQGTGTEIKPQNNPVPVSETNESSIVERT 2760  
Db 2701 QAKONVNGSVPMRTVGLNRLKNSFIQVADPDQGTGTEIKPQNNPVPVSETNESSIVERT 2760  
Qy 2761 PFSSSSSKHSSPSGTVAAARVTPFNYPNPSPRKSSADSTSARPSQIPTPVNNNTKKSDSKT 2820  
Db 2761 PFSSSSSKHSSPSGTVAAARVTPFNYPNPSPRKSSADSTSARPSQIPTPVNNNTKKSDSKT 2820  
Qy 2821 DSTESSGTQSPKHSGLSYLTVS 2843  
Db 2821 DSTESSGTQSPKHSGLSYLTVS 2843

RESULT 13  
US-08-289-548A-2  
; Sequence 2, Application US/08289548A  
; Patent No. 5648212  
; GENERAL INFORMATION:  
; APPLICANT: ALBERTSEN, HANS  
; APPLICANT: ANAND, RAKESH  
; APPLICANT: CARLSON, MARY  
; APPLICANT: GEODEN, JOANNA  
; APPLICANT: HEDGE, PHILIP J.  
; APPLICANT: JOSLYN, GEOFF  
; APPLICANT: KINZLER, KENNETH  
; APPLICANT: MARKHAM, ALEXANDER F.  
; APPLICANT: NAKAMURA, YUSUKE  
; APPLICANT: THLIVERIS, ANDREW  
; TITLE OF INVENTION: INHERITED AND SOMATIC MUTATIONS OF APC  
; TITLE OF INVENTION: GENE IN COLORECTAL CANCER IN HUMANS  
; NUMBER OF SEQUENCES: 102  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Banner & Allegretti, LTD  
; STREET: 1001 G Street, NW  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20001-4598  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA: US/08/289,548A  
; APPLICATION NUMBER: US/08/289,548A  
; FILING DATE: 12-AUG-1994  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Kagan, Sarah A.  
; REGISTRATION NUMBER: 32,141  
; REFERENCE/DOCKET NUMBER: 1107.46943  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-508-9100  
; TELEFAX: 202-508-9299  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2843 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-289-548A-2

Query Match

99.8%; Score 14533; DB 1; Length 2843;

Best Local Similarity 99.8%; Pred. No. 0;  
Matches 2836; Conservative 2; Mismatches 5; Indels 0; Gaps 0;  
Qy 1 MAAASYDQLLKQVZALKMENSINLRQELDNHNLTKLETEASNKMKVILKOLQGSIEDEAM 60  
Db 1 MAAASYDQLLKQVZALKMENSINLRQELDNHNLTKLETEASNKMKVILKOLQGSIEDEAM 60  
Qy 61 ASSGQIDLLERLKEINLDSSNPPGVKURSKMWSLRSYGRREGSVSSRSGECSPVPMGSPFR 120  
Db 61 ASSGQIDLLERLKEINLDSSNPPGVKURSKMWSLRSYGRREGSVSSRSGECSPVPMGSPFR 120  
Qy 121 RGFVNGSSESTGYLEELEKERSLLLLADLDKEEKMDWYIAOLNLTKEIDSLPTEFSL 180  
Db 121 RGFVNGSSESTGYLEELEKERSLLLLADLDKEEKMDWYIAOLNLTKEIDSLPTEFSL 180  
Qy 181 QTDLTRRQLEEARQIRVAMEEQLTGTCODMEKRAQRRIARIQQIEKOILRIRQLLSOAT 240  
Db 181 QTDLTRRQLEEARQIRVAMEEQLTGTCODMEKRAQRRIARIQQIEKOILRIRQLLSOAT 240  
Qy 241 EAERSSQNKHETGSHDAERQNEGGVGEINMATSNGCGSTTRVDHETASVLSSTHSA 300  
Db 241 EAERSSQNKHETGSHDAERQNEGGVGEINMATSNGCGSTTRVDHETASVLSSTHSA 300  
Qy 301 PRLTSLHGTKEVWYVYLLSMLGTHDKDDMSRTLLAMSSSDSCISMRQSGCLPLLIQLL 360  
Db 301 PRLTSLHGTKEVWYVYLLSMLGTHDKDDMSRTLLAMSSSDSCISMRQSGCLPLLIQLL 360  
Qy 361 HGNDKDSVLLGNRSGSKARARASAAHLNIIHSQDDKRGRRREIRVLHLEQIRAYCETC 420  
Db 361 HGNDKDSVLLGNRSGSKARARASAAHLNIIHSQDDKRGRRREIRVLHLEQIRAYCETC 420  
Qy 421 WEQEAHEPQMDQDNKMPAPVEHQICPAVCVLMKLSFDEHRAHNNELGLOAIBLQ 480  
Db 421 WEQEAHEPQMDQDNKMPAPVEHQICPAVCVLMKLSFDEHRAHNNELGLOAIBLQ 480  
Qy 481 VDCEMYGLTNDHYSITLRYAGMALTNLTFGDVANKATLCMKGCMRALVAQLKSEBDL 540  
Db 481 VDCEMYGLTNDHYSITLRYAGMALTNLTFGDVANKATLCMKGCMRALVAQLKSEBDL 540  
Qy 541 QOVVASVLRNLSRADVNSKKTLEVGSVKALMECALEVKESTLKSVLASLWNLASHCT 600  
Db 541 QOVVASVLRNLSRADVNSKKTLEVGSVKALMECALEVKESTLKSVLASLWNLASHCT 600  
Qy 601 ENKADICAVDGAFLVGTLYRSQTNLTALIESGGGILRVNSSLIATNEDHRIQIRENN 660  
Db 601 ENKADICAVDGAFLVGTLYRSQTNLTALIESGGGILRVNSSLIATNEDHRIQIRENN 660  
Qy 661 CLQTLLOHLKSHSLTIYSNACGTLWNLASRNPKDOEALWDMGAVSMLKNLIHSHKMIAM 720  
Db 661 CLQTLLOHLKSHSLTIYSNACGTLWNLASRNPKDOEALWDMGAVSMLKNLIHSHKMIAM 720  
Qy 721 GSAALRNLMANRPAKYKDANIMSPGSSLPVSHVRKQKALEAEALDAQHLSETFDNIDLS 780  
Db 721 GSAALRNLMANRPAKYKDANIMSPGSSLPVSHVRKQKALEAEALDAQHLSETFDNIDLS 780  
Qy 781 PKASHRSQKHQSLYGDYVDFTRHDDNRSDNFTGNMTVLSPLYLNTTVLPSSSSRGS 840  
Db 781 PKASHRSQKHQSLYGDYVDFTRHDDNRSDNFTGNMTVLSPLYLNTTVLPSSSSRGS 840  
Qy 841 LDSRSEKDRSLERERGLGNYHPATENPCTSKRGLOISTTAAQAKWMEVSAIHTS 900  
Db 841 LDSRSEKDRSLERERGLGNYHPATENPCTSKRGLOISTTAAQAKWMEVSAIHTS 900  
Qy 901 QEDRSSGTTTELHCVTDERNALRSSAAHTHSNTYNTFKSENSNRTCSMPYAKLEYKRSS 960  
Db 901 QEDRSSGTTTELHCVTDERNALRSSAAHTHSNTYNTFKSENSNRTCSMPYAKLEYKRSS 960  
Qy 961 NDSLNSVSSDGYKRGOMKPSIESYFDDDESFCISYGOVPADLAHKIHSANHMDNDGE 1020  
Db 961 NDSLNSVSSDGYKRGOMKPSIESYFDDDESFCISYGOVPADLAHKIHSANHMDNDGE 1020  
Qy 1021 LDTPINYSKYSDQLNSGRQSPSQNERWAPKHIIEDEIKQSPQRQSRNSTTYPVYTE 1080  
Db 1021 LDTPINYSKYSDQLNSGRQSPSQNERWAPKHIIEDEIKQSPQRQSRNSTTYPVYTE 1080

Db	1021	LDPTNYSLKYSDEGLNCSGROSPSONENMARPKHIEDEIKQSGORGRNOSTTYPYVTE	1080
Qy	1081	STDDHKLKFPQHPGQCEVSPYRSRSGANGSEPTNRVGNHGINOVSGQLCQEDDYEDDKP	1140
Db	1081	STDDHKLKFPQHPGQCEVSPYRSRSGANGSEPTNRVGNHGINOVSGQLCQEDDYEDDKP	1140
Qy	1141	TNYSRYSSEEOHHEERPTVYSIKYNEKRXHVOPIDYSLKXATDIPSSQKOSFSPRSKS	1200
Db	1141	TNYSRYSSEEOHHEERPTVYSIKYNEKRXHVOPIDYSLKXATDIPSSQKOSFSPRSKS	1200
Qy	1201	SSGQSKTEHMSSESTSTPSNAKRONQLHPSAQSRSQOPKATCKXSNINQETIQ	1260
Db	1201	SSGQSKTEHMSSESTSTPSNAKRONQLHPSAQSRSQOPKATCKXSNINQETIQ	1260
Qy	1261	TYCEDTPICFRCSGLSSLSABDEICNOTTOBADSANTLOJAEIKXGTSADBPV	1320
Db	1261	TYCEDTPICFRCSGLSSLSABDEICNOTTOBADSANTLOJAEIKXGTSADBPV	1320
Qy	1321	SEVPASQHPRTKSSRLQSSLSSESAHKAVERSSGAKSPKSGAOTPKSPZHYVQET	1380
Db	1321	SEVPASQHPRTKSSRLQSSLSSESAHKAVERSSGAKSPKSGAOTPKSPZHYVQET	1380
Qy	1381	PLMFSRCTSVSLDSFESRSIASSVQSPGCMVSGTISPSDLFDSPOQTMPPRSKTPP	1440
Db	1381	PLMFSRCTSVSLDSFESRSIASSVQSPGCMVSGTISPSDLFDSPOQTMPPRSKTPP	1440
Qy	1441	PPPOAOTKREVPKXKAFTAKRESGPXOAAVNAVORYVLPAADTLIHPATESTPDGF	1500
Db	1441	PPPOAOTKREVPKXKAFTAKRESGPXOAAVNAVORYVLPAADTLIHPATESTPDGF	1500
Qy	1501	SCSSLSALSJDEBFIOQVZLRIMPVQENDNGNETESBQPKESNENOEKAEKTIJSE	1560
Db	1501	SCSSLSALSJDEBFIOQVZLRIMPVQENDNGNETESBQPKESNENOEKAEKTIJSE	1560
Qy	1561	KDLDDSDDDDIIELEBCTIISAMPTKSSRKAQAQNAASKLPVPAVKPSPQLPYKLLPS	1620
Db	1561	KDLDDSDDDDIIELEBCTIISAMPTKSSRKAQAQNAASKLPVPAVKPSPQLPYKLLPS	1620
Qy	1621	QNRLOPQKGVHFTGDDKPRVYCVGEPINSTATISLSDLTIESPPMELAAAGEVIRGAQ	1680
Db	1621	QNRLOPQKGVHFTGDDKPRVYCVGEPINSTATISLSDLTIESPPMELAAAGEVIRGAQ	1680
Qy	1681	SGEPEKRDPTITBEGSTDEAOGKTSVTLPELDONKAEBGDIABECINSAMPKGSHPK	1740
Db	1681	SGEPEKRDPTITBEGSTDEAOGKTSVTLPELDONKAEBGDIABECINSAMPKGSHPK	1740
Qy	1741	FRVKKINDVVOOASASSAPKNOLDGKKKPTSPVXZIPONTYRTYVRKNASKNLN	1800
Db	1741	FRVKKINDVVOOASASSAPKNOLDGKKKPTSPVXZIPONTYRTYVRKNASKNLN	1800
Qy	1801	AERVFSNDKSKQNLKXNSKDFNDKLPNNEDRVRSFAPDSPHYTPIEGTPYCFGRND	1860
Db	1801	AERVFSNDKSKQNLKXNSKDFNDKLPNNEDRVRSFAPDSPHYTPIEGTPYCFGRND	1860
Qy	1861	SLSLSDPDDDDVLSREKAEIRKAKENKESAKTSHETLSNOOSANKTOAIKQPINR	1920
Db	1861	SLSLSDPDDDDVLSREKAEIRKAKENKESAKTSHETLSNOOSANKTOAIKQPINR	1920
Qy	1921	GQPKPILOKQSTFPOSSKDIIDRGATDEKQNFPAIENTPVCFSHNSLSLSIDIDENN	1980
Db	1921	GQPKPILOKQSTFPOSSKDIIDRGATDEKQNFPAIENTPVCFSHNSLSLSIDIDENN	1980
Qy	1981	NKEBPIETETPPPOGSPKPOASGAPKSPHVEDTPVCFSRNSSLSSIDEDDLQ	2040
Db	1981	NKEBPIETETPPPOGSPKPOASGAPKSPHVEDTPVCFSRNSSLSSIDEDDLQ	2040
Qy	2041	ECISSAMPKSKKPSRLKGDNEKSPRNMGILGELTLMDKDIORPSEHGLSDSENFD	2100
Db	2041	ECISSAMPKSKKPSRLKGDNEKSPRNMGILGELTLMDKDIORPSEHGLSDSENFD	2100
Qy	2101	WKAIOEGANSIVSSIHQAAAAACISROQSSDSITLSKSGISGSPHLLTPDOEKXPFT	2160
Db	2101	WKAIOEGANSIVSSIHQAAAAACISROQSSDSITLSKSGISGSPHLLTPDOEKXPFT	2160

  

Qy	2161	SNKGRILKPEXSTLETYKIESBKIGKGVKYSLLITGKVRNSBIISGMOKPLOAN	2220
Db	2161	SNKGRILKPEXSTLETYKIESBKIGKGVKYSLLITGKVRNSBIISGMOKPLOAN	2220
Qy	2221	MPISRGTMHHPGVNRNSSSTSPVSKKGPPLUKPASKSPSEGOATISPGKAXSVKS	2280
Db	2221	MPISRGTMHHPGVNRNSSSTSPVSKKGPPLUKPASKSPSEGOATISPGKAXSVKS	2280
Qy	2281	ELSFPVARTSOIGSSKADSPGSGRDSRSPAPQOLSPSPIOSPGRNSISPGRNGISPN	2340
Db	2281	ELSFPVARTSOIGSSKADSPGSGRDSRSPAPQOLSPSPIOSPGRNSISPGRNGISPN	2340
Qy	2341	KLSQLPRTSSPTASTKSSGSGMSTSPGRQMSQNLTKQGLSKMASIPRSBSASKG	2400
Db	2341	KLSQLPRTSSPTASTKSSGSGMSTSPGRQMSQNLTKQGLSKMASIPRSBSASKG	2400
Qy	2401	INOMNNGANKVLELRMSSTKSSGSESDRSERPVLVQSTFIKEAPSPTLRKLSESA	2460
Db	2401	INOMNNGANKVLELRMSSTKSSGSESDRSERPVLVQSTFIKEAPSPTLRKLSESA	2460
Qy	2461	SFESLSPPSRPASPTSOAQTPVLSPSLPDMLSHSSVQAGGMKLPNLSPTTEYNDG	2520
Db	2461	SFESLSPPSRPASPTSOAQTPVLSPSLPDMLSHSSVQAGGMKLPNLSPTTEYNDG	2520
Qy	2521	RPAKRHDIAHSSESPSRPLPNRSGTWKREHSHSSLPVRSWTMRRTGSSSSTLSASBS	2580
Db	2521	RPAKRHDIAHSSESPSRPLPNRSGTWKREHSHSSLPVRSWTMRRTGSSSSTLSASBS	2580
Qy	2581	SEKASEDEKHVNSISGTQSKENQVSAKGTWKIKENEFSPNTSQTIVSSGATNGAS	2640
Db	2581	SEKASEDEKHVNSISGTQSKENQVSAKGTWKIKENEFSPNTSQTIVSSGATNGAS	2640
Qy	2641	KTLIYQMAPAVSSTEDVWVRIEDCPINNRSGRSPGNTPPVIDSSEKAMPKIKOSKN	2700
Db	2641	KTLIYQMAPAVSSTEDVWVRIEDCPINNRSGRSPGNTPPVIDSSEKAMPKIKOSKN	2700
Qy	2701	OAKQVNGSVMPRTVGLNRLNLSFIQVADAPQKTEIRPGQNNPVPVSETNESSIVERT	2760
Db	2701	OAKQVNGSVMPRTVGLNRLNLSFIQVADAPQKTEIRPGQNNPVPVSETNESSIVERT	2760
Qy	2761	PFSSSSSSKHSBSSGTVAAKVPFNNBSPRKSADSTARSQJPTPVNNNTKKRDSKT	2820
Db	2761	PFSSSSSSKHSBSSGTVAAKVPFNNBSPRKSADSTARSQJPTPVNNNTKKRDSKT	2820
Qy	2821	DSTESSGTQSPKXHSGLVTSV	2843
Db	2821	DSTESSGTQSPKXHSGLVTSV	2843

RESULT 14  
 US-08-452-654-2  
 ; Sequence 2, Application US/08452654  
 ; Patent No. 5691454  
 ; GENERAL INFORMATION:  
 ; APPLICANT: ALBERTSEN, HANS  
 ; APPLICANT: ANAND, RAKESH  
 ; APPLICANT: CARLSON, MARY  
 ; APPLICANT: GRODEN, JOANNA  
 ; APPLICANT: HEDGE, PHILIP J.  
 ; APPLICANT: JOSLYN, GEOFF  
 ; APPLICANT: KINZLER, KENNETH  
 ; APPLICANT: MARKAK, ALEXANDER F.  
 ; APPLICANT: NAKAMURA, YUSUKE  
 ; APPLICANT: THLIVERIS, ANDREW  
 ; TITLE OF INVENTION: INHERITED AND SOMATIC MUTATIONS OF APC  
 ; NUMBER OF SEQUENCES: 94  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Hammer, Birch, McKie & Beckett  
 ; STREET: 1001 G Street, NW  
 ; CITY: Washington  
 ; STATE: D.C.

```

/ COUNTRY: USA
/ ZIP: 20001-4598
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Patent Release #1.0, Version #1.25
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/452,654
/ FILING DATE: 25-MAY-1995
/ CLASSIFICATION: 536
/ PRIORITY APPLICATION DATA:
/ APPLICATION NUMBER: US 07/741,940
/ FILING DATE: 08-AUG-1991
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Kagan, Sarah A.
/ REGISTRATION NUMBER: 32,141
/ REFERENCE/DOCKET NUMBER: 1107.035574
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 202-508-9100
/ TELEFAX: 202-508-9299
/ INFORMATION FOR SEQ ID NO: 2:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 2843 amino acids
/ TYPE: amino acid
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ US-08-452-654-2

Query Match 99.8%; Score 14533; DB 1; Length 2843;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2836; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 MAAASVDOLKQVEALKEKNSNLTROLEEDNSNHLTKLEFASNMKEVILKOLQGSIEDAM 60
DB 1 MAAASVDOLKQVEALKEKNSNLTROLEEDNSNHLTKLEFASNMKEVILKOLQGSIEDAM 60
QY 61 ASSGQIDLLERLKEINLDSSNFPGVKRSKMSLSYSGRSGSVSSRSGECSPVPMGSPFR 120
DB 61 ASSGQIDLLERLKEINLDSSNFPGVKRSKMSLSYSGRSGSVSSRSGECSPVPMGSPFR 120
QY 61 ASSGQIDLLERLKEINLDSSNFPGVKRSKMSLSYSGRSGSVSSRSGECSPVPMGSPFR 120
DB 61 ASSGQIDLLERLKEINLDSSNFPGVKRSKMSLSYSGRSGSVSSRSGECSPVPMGSPFR 120
QY 121 RGVNVSRESSTGYLELEKEKRSLLADLDKEKEKDWYAAQQLTKRIDLPLETENSL 180
DB 121 RGVNVSRESSTGYLELEKEKRSLLADLDKEKEKDWYAAQQLTKRIDLPLETENSL 180
QY 121 RGVNVSRESSTGYLELEKEKRSLLADLDKEKEKDWYAAQQLTKRIDLPLETENSL 180
DB 121 RGVNVSRESSTGYLELEKEKRSLLADLDKEKEKDWYAAQQLTKRIDLPLETENSL 180
QY 181 QTDMTNRQLEYEARQIRVAMEQOLGTCODEMKAQRRIARIQIIEKIDILRIROLLQSGAT 240
DB 181 QTDMTNRQLEYEARQIRVAMEQOLGTCODEMKAQRRIARIQIIEKIDILRIROLLQSGAT 240
QY 181 QTDMTNRQLEYEARQIRVAMEQOLGTCODEMKAQRRIARIQIIEKIDILRIROLLQSGAT 240
DB 181 QTDMTNRQLEYEARQIRVAMEQOLGTCODEMKAQRRIARIQIIEKIDILRIROLLQSGAT 240
QY 241 EAEBSQNKERTGSHDAERQNGQVGEINMATSNGQGSTTRMDHETASVLSSTHSA 300
DB 241 EAEBSQNKERTGSHDAERQNGQVGEINMATSNGQGSTTRMDHETASVLSSTHSA 300
QY 241 EAEBSQNKERTGSHDAERQNGQVGEINMATSNGQGSTTRMDHETASVLSSTHSA 300
DB 241 EAEBSQNKERTGSHDAERQNGQVGEINMATSNGQGSTTRMDHETASVLSSTHSA 300
QY 301 PARLISHLGTKEYMYSLISM.LGTHDKDMSTRLLAMSSQDSISMRSQSGCLPLLIQL 360
DB 301 PARLISHLGTKEYMYSLISM.LGTHDKDMSTRLLAMSSQDSISMRSQSGCLPLLIQL 360
QY 301 PARLISHLGTKEYMYSLISM.LGTHDKDMSTRLLAMSSQDSISMRSQSGCLPLLIQL 360
DB 301 PARLISHLGTKEYMYSLISM.LGTHDKDMSTRLLAMSSQDSISMRSQSGCLPLLIQL 360
QY 361 HGNDDXSVLLGNSRSGKEAPAPASALAHNIHSQPDYKRGREIRFVLLHLEQIRAYCETC 420
DB 361 HGNDDXSVLLGNSRSGKEAPAPASALAHNIHSQPDYKRGREIRFVLLHLEQIRAYCETC 420
QY 361 HGNDDXSVLLGNSRSGKEAPAPASALAHNIHSQPDYKRGREIRFVLLHLEQIRAYCETC 420
DB 361 HGNDDXSVLLGNSRSGKEAPAPASALAHNIHSQPDYKRGREIRFVLLHLEQIRAYCETC 420
QY 421 MEMOAHBEHGMQDKNMPAPVEHQICPAVCYLM.LSPDEHRHAMNELGGLQAIAYELIQ 480
DB 421 MEMOAHBEHGMQDKNMPAPVEHQICPAVCYLM.LSPDEHRHAMNELGGLQAIAYELIQ 480
QY 421 MEMOAHBEHGMQDKNMPAPVEHQICPAVCYLM.LSPDEHRHAMNELGGLQAIAYELIQ 480
DB 421 MEMOAHBEHGMQDKNMPAPVEHQICPAVCYLM.LSPDEHRHAMNELGGLQAIAYELIQ 480
QY 481 VDCENYGLTNDHYSITLRRYAGMALTNLTFGDVANKATLCSNKGCMRALVAQLKSESEDL 540
DB 481 VDCENYGLTNDHYSITLRRYAGMALTNLTFGDVANKATLCSNKGCMRALVAQLKSESEDL 540
QY 481 VDCENYGLTNDHYSITLRRYAGMALTNLTFGDVANKATLCSNKGCMRALVAQLKSESEDL 540
DB 481 VDCENYGLTNDHYSITLRRYAGMALTNLTFGDVANKATLCSNKGCMRALVAQLKSESEDL 540
QY 541 QOVIVASVLRNLISWRADVNSKTLREVGSKALMEGALVYKKESTLKYLSALMINSACT 600
DB 541 QOVIVASVLRNLISWRADVNSKTLREVGSKALMEGALVYKKESTLKYLSALMINSACT 600
QY 541 QOVIVASVLRNLISWRADVNSKTLREVGSKALMEGALVYKKESTLKYLSALMINSACT 600
DB 541 QOVIVASVLRNLISWRADVNSKTLREVGSKALMEGALVYKKESTLKYLSALMINSACT 600

```

```

QY 601 ENKADIICAVDGAALFVGLTYRSQNTLAIIESGGGILIRNYSLLATNEDEHROILRENN 660
DB 601 ENKADIICAVDGAALFVGLTYRSQNTLAIIESGGGILIRNYSLLATNEDEHROILRENN 660
QY 661 CLQTLLOHLKSHSLITVSNACGLTMINLSARNPDQDALMDGAVSNLKNLISHKEMIAM 720
DB 661 CLQTLLOHLKSHSLITVSNACGLTMINLSARNPDQDALMDGAVSNLKNLISHKEMIAM 720
QY 721 GSAALRLNMANPACTYKANTMSPESSILPSLHYRKQKLELDAQHISEFPDNDINS 780
DB 721 GSAALRLNMANPACTYKANTMSPESSILPSLHYRKQKLELDAQHISEFPDNDINS 780
QY 781 PKASHRSKQKHQSLGYVFDINRHDNRSDNFTGNMTVASPYLNTLVLPSSSSRSGS 840
DB 781 PKASHRSKQKHQSLGYVFDINRHDNRSDNFTGNMTVASPYLNTLVLPSSSSRSGS 840
QY 841 LDSRSSEKORSLEBERGIGLGNYPATENPGISKRGLQISTTAQIAYMEVSAIHS 900
DB 841 LDSRSSEKORSLEBERGIGLGNYPATENPGISKRGLQISTTAQIAYMEVSAIHS 900
QY 901 QEDRSSGSTTEHLCVDERNALRSSAAHTSTNTYFTSENSNRCTSPYAKLEYKRS 960
DB 901 QEDRSSGSTTEHLCVDERNALRSSAAHTSTNTYFTSENSNRCTSPYAKLEYKRS 960
QY 961 NDSLSVSSSDGYGKRGQMKPSIESYEDDESKFCGYPADLAHTISANHMNDNCE 1020
DB 961 NDSLSVSSSDGYGKRGQMKPSIESYEDDESKFCGYPADLAHTISANHMNDNCE 1020
QY 1021 LDPPIVSLKYSDQELNSGRQSPQNERMARPHIIEDEIKQSEORQSNQSTTYPVYTE 1080
DB 1021 LDPPIVSLKYSDQELNSGRQSPQNERMARPHIIEDEIKQSEORQSNQSTTYPVYTE 1080
QY 1081 STDQKHLKQPHFGQOEVSYPYRSRGANGSETRNVSGNIGNVNSQSLCOEDDYEDDXP 1140
DB 1081 STDQKHLKQPHFGQOEVSYPYRSRGANGSETRNVSGNIGNVNSQSLCOEDDYEDDXP 1140
QY 1141 TNGERYSEEEQEEBERPTNYSIKNEKRAHYDQIDYSLKANDIPSSQOSRFSFS 1200
DB 1141 TNGERYSEEEQEEBERPTNYSIKNEKRAHYDQIDYSLKANDIPSSQOSRFSFS 1200
QY 1201 SSGQSKTEHMSSSSNTSTPSSNAKRONQLHPSASQSRGQPOKATCKVSSINQETIQ 1260
DB 1201 SSGQSKTEHMSSSSNTSTPSSNAKRONQLHPSASQSRGQPOKATCKVSSINQETIQ 1260
QY 1261 TYCEVEDTPICFRSCSSLSLSAEDBIGCQNTTOEADSANTLOIANIKKIGTRSAEDV 1320
DB 1261 TYCEVEDTPICFRSCSSLSLSAEDBIGCQNTTOEADSANTLOIANIKKIGTRSAEDV 1320
QY 1321 SEVPAYSOHPRTKSRLOQSSLSSEBARHKAVEFFSGAKSPKSGAQTPKSPPEHYQET 1380
DB 1321 SEVPAYSOHPRTKSRLOQSSLSSEBARHKAVEFFSGAKSPKSGAQTPKSPPEHYQET 1380
QY 1381 PLMFSRCTSVSLDSFESRSIASSVQSEPCGWNVSGIISPDI.PDSFGQTMPPSRKTPP 1440
DB 1381 PLMFSRCTSVSLDSFESRSIASSVQSEPCGWNVSGIISPDI.PDSFGQTMPPSRKTPP 1440
QY 1441 PPQOTQKREVKNAFPAEKRESGPKQAAVAAVQVQVLPDADTLHPFTESIPDGF 1500
DB 1441 PPQOTQKREVKNAFPAEKRESGPKQAAVAAVQVQVLPDADTLHPFTESIPDGF 1500
QY 1501 SCSSLSALSIDPEPIQKQVEALRIMPVOENDNGNTESEOPKESNENQEKAEKTIQSE 1560
DB 1501 SCSSLSALSIDPEPIQKQVEALRIMPVOENDNGNTESEOPKESNENQEKAEKTIQSE 1560
QY 1561 KDLIDSDDDDDIIEECIISAMPYTSRKAAPKPAOTAKLPPVARKSQOLPVYTLPS 1620
DB 1561 KDLIDSDDDDDIIEECIISAMPYTSRKAAPKPAOTAKLPPVARKSQOLPVYTLPS 1620
QY 1621 QNRLOPQKHSFTPGDMFRVYVECTPINFSTATSLDLTIESPNEILAGBGVAGGAQ 1680
DB 1621 QNRLOPQKHSFTPGDMFRVYVECTPINFSTATSLDLTIESPNEILAGBGVAGGAQ 1680
QY 1681 SGEFEKRDITIPTEGASTDEAQGGKISSVTIIPBLDNKABEGDILAEICINSAPKKGSHKP 1740

```

```

Db 1681 SCEPEKNTITTEGRSTDEAGCGTSSATTIPELDNDNAEBDILAEELNAMPKSKHNP 1740
Qy 1741 FRVKKINDQVOOASASSAPNNKOLDGKKKPTSPVKRIPONTYRFRVRKANADSKNIN 1800
Db 1741 FRVKKINDQVOOASASSAPNNKOLDGKKKPTSPVKRIPONTYRFRVRKANADSKNIN 1800
Qy 1801 AERYVSDNDKSKONLKNKSKDFNDKLPNNEDVRGSAFSPHHTPIGTPICFGRND 1860
Db 1801 AERYVSDNDKSKONLKNKSKDFNDKLPNNEDVRGSAFSPHHTPIGTPICFGRND 1860
Qy 1861 SLSSLDPDDDDVDLSREKALRKAKENKESBAKTSHTLETSSNOQANKTOALAKOPINR 1920
Db 1861 SLSSLDPDDDDVDLSREKALRKAKENKESBAKTSHTLETSSNOQANKTOALAKOPINR 1920
Qy 1921 GQKPIIQQKSTFPOSSKDIDRGAATDEKIQNTAIENTPVCFSHNSLSLSDIDENN 1980
Db 1921 GQKPIIQQKSTFPOSSKDIDRGAATDEKIQNTAIENTPVCFSHNSLSLSDIDENN 1980
Qy 1981 NKENEPIKETEPPOGEPSPKQASGYAPKSFHYEDTPVCFGRNSLSLSDIDEDLLQ 2040
Db 1981 NKENEPIKETEPPOGEPSPKQASGYAPKSFHYEDTPVCFGRNSLSLSDIDEDLLQ 2040
Qy 2041 ECISSNAPKKKKPSRLKGDNEKSPRNWGGILGEDLTLDKIDRPPSEHGLSPDSNFD 2100
Db 2041 ECISSNAPKKKKPSRLKGDNEKSPRNWGGILGEDLTLDKIDRPPSEHGLSPDSNFD 2100
Qy 2101 WKALQEGANSIVSSHQAACLSROASDSDSLSLKSGISLSPFHLTPDOEKXPT 2160
Db 2101 WKALQEGANSIVSSHQAACLSROASDSDSLSLKSGISLSPFHLTPDOEKXPT 2160
Qy 2161 SNKSPRIILKPEKSTLETKLIESKSGIKGKRYKSLITGVANSSEISGOMKQLOAN 2220
Db 2161 SNKSPRIILKPEKSTLETKLIESKSGIKGKRYKSLITGVANSSEISGOMKQLOAN 2220
Qy 2221 MPSTSRGTMTHIGVNRSSSTSPVSKKGBPLTPASKSBEQCTATTSPRGAKEPVKS 2280
Db 2221 MPSTSRGTMTHIGVNRSSSTSPVSKKGBPLTPASKSBEQCTATTSPRGAKEPVKS 2280
Qy 2281 ELSPVARQTSQIGSSSKAPSRSGRSDTSPSRPAQPSRPIQSPGRNSISPRGANGISPPN 2340
Db 2281 ELSPVARQTSQIGSSSKAPSRSGRSDTSPSRPAQPSRPIQSPGRNSISPRGANGISPPN 2340
Qy 2341 KLSQUPRTSSPSTASTKSSGSKMSYSPRGOMQONLTQKTGSKNAASSIPRESASKG 2400
Db 2341 KLSQUPRTSSPSTASTKSSGSKMSYSPRGOMQONLTQKTGSKNAASSIPRESASKG 2400
Qy 2401 LNOMNNGANGAKKYTELSRMSSTKSSGSRSDRSPVAVROSTFIKEAPSPILRKLESA 2460
Db 2401 LNOMNNGANGAKKYTELSRMSSTKSSGSRSDRSPVAVROSTFIKEAPSPILRKLESA 2460
Qy 2461 SFESLSPPSPASPTPSQAQTPVLSPLPDMSLSTHSSVQAGMRKLPENISPTIEVNDG 2520
Db 2461 SFESLSPPSPASPTPSQAQTPVLSPLPDMSLSTHSSVQAGMRKLPENISPTIEVNDG 2520
Qy 2521 RPARKHDIDARSHSBSRPLPINRSGTMRKSHSSSLPRVSTMRKRGSSSSILASSES 2580
Db 2521 RPARKHDIDARSHSBSRPLPINRSGTMRKSHSSSLPRVSTMRKRGSSSSILASSES 2580
Qy 2581 SEKASSEDEKAVNISGTQKQENOVSAKGTWRKIKENEFSPITNSQTVSSGATNGAES 2640
Db 2581 SEKASSEDEKAVNISGTQKQENOVSAKGTWRKIKENEFSPITNSQTVSSGATNGAES 2640
Qy 2641 KTLIYQMAPAVSKTEDVWRLEDCEPINRSGRSPGTGTPVIDSVSKAPNPKDSDN 2700
Db 2641 KTLIYQMAPAVSKTEDVWRLEDCEPINRSGRSPGTGTPVIDSVSKAPNPKDSDN 2700
Qy 2701 QAKQNVGSGVPMRTVGLLENLSFIQVADAPDOKTEIKPGONNPVVPSETNESSIYERT 2760
Db 2701 QAKQNVGSGVPMRTVGLLENLSFIQVADAPDOKTEIKPGONNPVVPSETNESSIYERT 2760
Qy 2761 PPSSSSSKSHSGSPGTVAARVTPFVYNPSPKSSADSTISAPSQIPTPVNNNTKKRDSKT 2820

```

```

Db 2761 PPSSSSSKSHSGSPGTVAARVTPFVYNPSPKSSADSTISAPSQIPTPVNNNTKKRDSKT 2820
Qy 2821 DSTESSGTQSPKRHSGSYLVTSV 2843
Db 2821 DSTESSGTQSPKRHSGSYLVTSV 2843

RESULT 15
US-08-370-235A-2
; Sequence 2, Application US/08370235A
; Patent No. 5910418
; GENERAL INFORMATION:
; APPLICANT: VOGELSTEIN, BERT
; APPLICANT: KINZLER, KENNETH W.
; APPLICANT: HILL, DAVID E.
; APPLICANT: JOHNSON, KAREN A.
; TITLE OF INVENTION: ANTIBODIES AND ASSAYS FOR DETERMINING
; TITLE OF INVENTION: MUTATIONS IN THE APC GENE
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: BANNER & WITCOFF, LTD.
; STREET: 1001 G STREET, N.W.
; CITY: WASHINGTON
; STATE: DC
; COUNTRY: US
; ZIP: 20001
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/370,235A
; FILING DATE: 01-JAN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: KAGAN, SARAH A.
; REGISTRATION NUMBER: 32,141
; REFERENCE/DOCKET NUMBER: 01107,48668
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202 508 9299
; TELEFAX: 202 508 9299
; INFORMATION FOR SEQ. ID NO.: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2843 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-370-235A-2

Query Match 99.8%; Score 14533; DB 2; Length 2843;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2836; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

1 MAASVQDILLERLKEKLNLDSSNPPGVKLRKMSIRSYSGSEGSVSRSGCSPVPMGSPFR 120
1 AASSGQIDILLERLKEKLNLDSSNPPGVKLRKMSIRSYSGSEGSVSRSGCSPVPMGSPFR 120
61 AASSGQIDILLERLKEKLNLDSSNPPGVKLRKMSIRSYSGSEGSVSRSGCSPVPMGSPFR 120
61 AASSGQIDILLERLKEKLNLDSSNPPGVKLRKMSIRSYSGSEGSVSRSGCSPVPMGSPFR 120
121 RGFVNSRSTGTGLELEKERSLLADLDKPEKEXOWYAAQLONLKRIDSLPTENFSL 180
121 RGFVNSRSTGTGLELEKERSLLADLDKPEKEXOWYAAQLONLKRIDSLPTENFSL 180
181 QTDMTROLLEYEARQIVAMEEOLGTQDMERAKORRIARIQIEKDIRIRQLQSQAT 240
181 QTDMTROLLEYEARQIVAMEEOLGTQDMERAKORRIARIQIEKDIRIRQLQSQAT 240
241 EARRSSONHGHGSHARONGGGVGEINMATSGGCGSTTRMDETAASYSSSTGSA 300
241 EARRSSONHGHGSHARONGGGVGEINMATSGGCGSTTRMDETAASYSSSTGSA 300

```



QY 301 PRLTSHLGTKEWYIILLMLGTHDKODMSRTLLAMSSSQDSICSMQSGCLPILLIQL 360  
 Db 301 PRLTSHLGTKEWYIILLMLGTHDKODMSRTLLAMSSSQDSICSMQSGCLPILLIQL 360  
 QY 361 HENDSDVTLNRSRSGKEARASALNIIHSDDDDRGREIRVUILLIOIAYCETC 420  
 Db 361 HENDSDVTLNRSRSGKEARASALNIIHSDDDDRGREIRVUILLIOIAYCETC 420  
 QY 421 MEMOEAEHSDGMDQXNPMAPVEHOICPAVCYMLLSPDEEHRANMELGLOIAIELLO 480  
 Db 421 MEMOEAEHSDGMDQXNPMAPVEHOICPAVCYMLLSPDEEHRANMELGLOIAIELLO 480  
 QY 481 VDCENYGLTNDHYSTTLRRYAGMALTNLTFGDVANKATLCISMGKMLVUQLKSEEDL 540  
 Db 481 VDCENYGLTNDHYSTTLRRYAGMALTNLTFGDVANKATLCISMGKMLVUQLKSEEDL 540  
 QY 541 QCVIASVJNLMSRADVNSKTLREVGSYKALMECALVEYKKESTLKSVALMNSACT 600  
 Db 541 QCVIASVJNLMSRADVNSKTLREVGSYKALMECALVEYKKESTLKSVALMNSACT 600  
 QY 601 ENKADI CAVDGALALVGLTYRSGOTNTLAIIESGGILRNVSJLIATNEHROILRENN 660  
 Db 601 ENKADI CAVDGALALVGLTYRSGOTNTLAIIESGGILRNVSJLIATNEHROILRENN 660  
 QY 661 CLOTLLOHLKSHSLCTIVSACGLTWNLSARNPKOQALMDMGAVSMLJHSKXMTAM 720  
 Db 661 CLOTLLOHLKSHSLCTIVSACGLTWNLSARNPKOQALMDMGAVSMLJHSKXMTAM 720  
 QY 721 GSAALRLNIMANRPAYKODANTMSPGSLPSLHVKKQALAEALDACHLSETPNDLNS 780  
 Db 721 GSAALRLNIMANRPAYKODANTMSPGSLPSLHVKKQALAEALDACHLSETPNDLNS 780  
 QY 781 PASHRSKORHKOSLYGDVYPTNRHDDNSDNTGNTVLSPLYNTTVLPSSSSSGS 840  
 Db 781 PASHRSKORHKOSLYGDVYPTNRHDDNSDNTGNTVLSPLYNTTVLPSSSSSGS 840  
 QY 841 LPSRSRSDKRLSERERIGIGLANYHAPATENPGTSSKRLQISTTAQAIKXMEVSAIHTS 900  
 Db 841 LPSRSRSDKRLSERERIGIGLANYHAPATENPGTSSKRLQISTTAQAIKXMEVSAIHTS 900  
 QY 901 QEDRSRSGTTEHACTTDERNARSSAAHTSNTNFTKSEKSNFTCMFPAKLEFKSS 960  
 Db 901 QEDRSRSGTTEHACTTDERNARSSAAHTSNTNFTKSEKSNFTCMFPAKLEFKSS 960  
 QY 961 NDLNLSVSSSDGYGRGQMKPSIESYSEDDSKFCSYGOYPADLAKHISANEMDNDGE 1020  
 Db 961 NDLNLSVSSSDGYGRGQMKPSIESYSEDDSKFCSYGOYPADLAKHISANEMDNDGE 1020  
 QY 1021 LDTPTNYSIKYSDEQLNSGRSPSONERWAPKXIIIDELIKOSEGROGRNOSTIYVYTE 1080  
 Db 1021 LDTPTNYSIKYSDEQLNSGRSPSONERWAPKXIIIDELIKOSEGROGRNOSTIYVYTE 1080  
 QY 1081 STDDHLKFOHPHGOECVSPYRSGANGSEFNRGSHNGINQVNSQSLCEDEDIYEDKP 1140  
 Db 1081 STDDHLKFOHPHGOECVSPYRSGANGSEFNRGSHNGINQVNSQSLCEDEDIYEDKP 1140  
 QY 1141 TNYSERYSBEEQHEEERPTNYSIKYNEKRAHVDPIDYSLKYATDIPSSQKOSFSGS 1200  
 Db 1141 TNYSERYSBEEQHEEERPTNYSIKYNEKRAHVDPIDYSLKYATDIPSSQKOSFSGS 1200  
 QY 1201 SSGGSKTEHMSSESTSTPSNAKRONOLHPSAOSRSGOPQAACTCKVSSINQETIQ 1260  
 Db 1201 SSGGSKTEHMSSESTSTPSNAKRONOLHPSAOSRSGOPQAACTCKVSSINQETIQ 1260  
 QY 1261 TVCVDDTPICFRCSLSLSLSAEDEICGNOTTOADASANTLQIAELIKGITSABEDV 1320  
 Db 1261 TVCVDDTPICFRCSLSLSLSAEDEICGNOTTOADASANTLQIAELIKGITSABEDV 1320  
 QY 1321 SEVPAVSQHPRTKSSRLQSSLSSESAAHKAVERSGAKSPKSAQOTPKSPPEHYOET 1380  
 Db 1321 SEVPAVSQHPRTKSSRLQSSLSSESAAHKAVERSGAKSPKSAQOTPKSPPEHYOET 1380

QY 1381 PLMPSRCTVSLSLSDTESRSIASSVQSPQSGWVSGIIEPSSLPDSPGQTPMPSRKTTP 1440  
 Db 1381 PLMPSRCTVSLSLSDTESRSIASSVQSPQSGWVSGIIEPSSLPDSPGQTPMPSRKTTP 1440  
 QY 1441 PPOQTAKREVPKXNAPFAEKRESSEGPQOAVAAQVQVLPDADTLHFAFTESTPDGF 1500  
 Db 1441 PPOQTAKREVPKXNAPFAEKRESSEGPQOAVAAQVQVLPDADTLHFAFTESTPDGF 1500  
 QY 1501 SCSSLSLALSDBPFIQXVLEIRIMPVOENDNGNTESEOPKESNENOEKAEKTI0SE 1560  
 Db 1501 SCSSLSLALSDBPFIQXVLEIRIMPVOENDNGNTESEOPKESNENOEKAEKTI0SE 1560  
 QY 1561 KDLILDSDDDDIIELECTISAMPKTSRKAAPKAPATAKLPVPVARKSQLPVYKILPS 1620  
 Db 1561 KDLILDSDDDDIIELECTISAMPKTSRKAAPKAPATAKLPVPVARKSQLPVYKILPS 1620  
 QY 1621 QNLQOPKHVSFTPGDMPRVYCEGTPINFSTATISLSDLTIESPPNELAAGEVGAQ 1680  
 Db 1621 QNLQOPKHVSFTPGDMPRVYCEGTPINFSTATISLSDLTIESPPNELAAGEVGAQ 1680  
 QY 1681 SGSEFERDITPTEGRSTDAQGGKTSVITPELDONKAEEOGILAECLNSAMPKGSHP 1740  
 Db 1681 SGSEFERDITPTEGRSTDAQGGKTSVITPELDONKAEEOGILAECLNSAMPKGSHP 1740  
 QY 1741 FRVKKIMDQVOQASASSAPNKQOLDGKKKQTSVPKP1PONTYRTRVRXNADSKNILN 1800  
 Db 1741 FRVKKIMDQVOQASASSAPNKQOLDGKKKQTSVPKP1PONTYRTRVRXNADSKNILN 1800  
 QY 1801 AERVFSDNDSKQKONKNNKDFNDLPNNEDVRSSFAFSDPHHTPIEGFPYCFSRD 1860  
 Db 1801 AERVFSDNDSKQKONKNNKDFNDLPNNEDVRSSFAFSDPHHTPIEGFPYCFSRD 1860  
 QY 1861 SLSSLDPDDDDVLSREKALERAKENKSEAKVTSHTLTIQNOASANKTOAIKOPINR 1920  
 Db 1861 SLSSLDPDDDDVLSREKALERAKENKSEAKVTSHTLTIQNOASANKTOAIKOPINR 1920  
 QY 1921 GQCPKILQKOSTPQSSKDI0PRGAATDEKONPAIENTPVCFSHNSLSLSLSDIDQEN 1980  
 Db 1921 GQCPKILQKOSTPQSSKDI0PRGAATDEKONPAIENTPVCFSHNSLSLSLSDIDQEN 1980  
 QY 1981 NKENEPKETEPPDQSGEBSKQASGAPKSFHVEDTPVCFSRNSLSLSLSDSEDDLIQ 2040  
 Db 1981 NKENEPKETEPPDQSGEBSKQASGAPKSFHVEDTPVCFSRNSLSLSLSDSEDDLIQ 2040  
 QY 2041 ECISAMPKXKXSRILKGNKESPRNMGGIIGEDTLTDLKDIORPDSHGLSPSENF 2100  
 Db 2041 ECISAMPKXKXSRILKGNKESPRNMGGIIGEDTLTDLKDIORPDSHGLSPSENF 2100  
 QY 2101 WKALQEGANSIVSLQAAAAACLSRQASDSDSILSKSGISLQSPFHLTPDQEKPT 2160  
 Db 2101 WKALQEGANSIVSLQAAAAACLSRQASDSDSILSKSGISLQSPFHLTPDQEKPT 2160  
 QY 2161 SNKGPRILKPGKSTLETKKIESESGKIGKQKVVYSLITGVRNSNBEISGQKOPLOAN 2220  
 Db 2161 SNKGPRILKPGKSTLETKKIESESGKIGKQKVVYSLITGVRNSNBEISGQKOPLOAN 2220  
 QY 2221 MBSISGRMTMHTPGVRNSSSTSPVSKGQPLKTPASBSBEGOTATISPRGAAPSVYS 2280  
 Db 2221 MBSISGRMTMHTPGVRNSSSTSPVSKGQPLKTPASBSBEGOTATISPRGAAPSVYS 2280  
 QY 2281 ELSPFVARQTSQIGSSKAPSRSGSRDSTPSRAQOPLSPPIOSPGNNSISPRNGISPPN 2340  
 Db 2281 ELSPFVARQTSQIGSSKAPSRSGSRDSTPSRAQOPLSPPIOSPGNNSISPRNGISPPN 2340  
 QY 2341 KLSQLPRTSSPTASTKSSGSGKMYTSPGRMOQONLTQGLSKNASSIPRSSASKG 2400  
 Db 2341 KLSQLPRTSSPTASTKSSGSGKMYTSPGRMOQONLTQGLSKNASSIPRSSASKG 2400  
 QY 2401 LNMNNNGANKKVELSRWSSTKSSGSESDBRBRPVLVQSFIFIEAPSPTLRKLBSBA 2460  
 Db 2401 LNMNNNGANKKVELSRWSSTKSSGSESDBRBRPVLVQSFIFIEAPSPTLRKLBSBA 2460  
 QY 2461 SFESLSFSSRPASPTRSQAQTPVLSPLSDMSLSTHSSVQAGWRLPNNLSPTEIYNDG 2520





**This Page Blank (uspio)**